Measurement of the top pair-production in association with a Z boson at LHC and WZ scattering in view of the High Luminosity LHC

Deniz Poyraz

CERN-THESIS-2018-125
19/06/2018

Supervisors: Prof. Dr. Didar Dobur, Dr. Michael Tytgat

Proefschrift tot het bekomen van de graad van
Doctor in de Wetenschappen: Fysica
Academiejaar 2017-2018

Universiteit Gent
Faculteit Wetenschappen
Vakgroep Fysica en Sterrenkunde

Members of the examination committee

dr. Wim Cosyn (University of Ghent)
prof. dr. Christophe Detavernier (University of Ghent)
prof. dr. Natalie Jachowicz (University of Ghent)
prof. dr. Dirk Ryckbosch (University of Ghent)
dr. Pedro Ferreira da Silva (CERN)
dr. Efe Yazgan (Institute of High Energy Physics of the Chinese Academy of Sciences)

You can know the name of that bird in all the languages of the world, but when you're finished, you'll know absolutely nothing whatever about the bird. You'll only know about humans in different places, and what they call the bird...
I learned very early the difference between knowing the name of something and knowing something.

Richard P. Feynman

Acknowledgements

This thesis is the outcome of four years and nine months in which I had a joyful experience in a great scientific environment and met with many people that without them, this journey would not be the same.

First and foremost I would like to thank to Michael for giving me the opportunity to start this position and for always being there. Dear Dirk, you are one of the kindest persons I have met during these years. Thank you for having your door always open to me and also for your advices.

Efe, you helped me a lot with all the new things in CMS in my first years. I thank you for your guidance, for useful discussions and also for introducing me to the upgrade study. Apart from work, it was nice to have your and Yasemin's friendship during my Cern visits. I also like to thank Bugra for the good times at Cern.

I would like to express my great attitude to the people I worked with for the upgrade study in my second year. I have learned a lot in this four months of stay at Cern. Pietro Govoni, Raffaele Gerosa, Xavier Janssen, Jasper and Efe, it was a pleasure to work with you all.

As I move to my third year in time, I would like to deeply thank to Didar. I must admit some days were hard and challenging with all the deadlines we try to make. But at the end, those moments made me learn a lot and also a stronger person with a different perspective. I can not count how many times I had knocked your door for what different reasons, you always had time for me with a positive attitude. It was great to have your courage and support. I specifically like to thank you for the very late night we went through in Skype during Moriond, this is just one of the many moments I will remember with a smile afterwards. Thank you for all the discussions and work we have done together and also for the nice conversations.

I would like to thank to friendly Ghent - CMS people. Illia, I sometimes can not believe how we had survived all the deadlines and the synchronisations. Thank you for all the discussions and your collaboration. Robert, Tom, Willem, Daniele, Martina thank you guys for all the discussions,chats and also lunch-coffee breaks. Special thanks to Tom for reading and correcting my thesis. Additionally I like to thank to CMS-Top group, especially to Mirena, Andrea, Pedro and Rebeca.

The best office mates ever, Celine and Ianthe. I am very glad that I moved to your office. You helped me to cheer up in some difficult days and also share my happiness always. It was very pleasant to share the same office with you, thank you. I would like to thank to my colleagues in INW for lunch-coffee breaks and for the after work activities, especially to Sam and Sander. Sinem meeting you in INW when I first arrived, was such a surprise

for me. I can not imagine how life would be here otherwise. Martina, I am so happy for knowing you before I finish my PhD. I am very grateful to your positive attitude that certainly made my days easier and more joyful during the writing period. All the times I had with you girls in INW, at Cern (pity we were not at Cern together with you Martina at the same times) and in many other circumstances were all unforgettable, Volare Trieste! I also like to thank to my friends Gozde and Ece for the great times in Belgium and in Germany and to Deniz for being with me always even from far away. The last chapters of this thesis were written in University of Antwerp, I like to thank to Nick van Remortel for this and to all colleagues for the nice lunch breaks.

Of course my grateful thanks go to my parents. I thank them for supporting and encouraging me in every decision I make. I deeply thank to my aunt Semra for being there all the times I need and also for taking care of me in Idstein after my deadlines and to my aunt Berna for the joyful times we had during Christmas holidays. My deep love and thanks go to my grandmother. She also gave me the opportunity to write some part of my thesis in her house by the sea which was an amazing chance. Here would be the right moment to also thank my grandfather Burhan Cahit Unal for giving me the first inspiration when I was in high school. All my journey had started thanks to you, you were my idol in physics and also someways in life. I thank you for asking me many questions about how and why things happen in every circumstances when I was a kid. I am very proud of having a similar path as yours. I like to thank to Yasmine for always being a great company and friend, especially during the times we shared the same house together. Zo blij met iemand als jij. Caroline and Rudi, bedankt voor alles.

Last but not least, Jasper meeting with you was the best thing happened to me in these years. The moments and the days we spent together were all very precious. It was also great chance that I shared the same office with you at Cern, worked on the same project, took shifts at P5 and many more. Thank you for being with me always. Now I look forward to share many more happy days with you.

Abstract

This thesis describes two studies towards measurements of couplings of heavy particles of the standard model that will be useful in establishing the electroweak symmetry breaking. The first study describes a cross section measurement of top quark pair production in association with a Z boson ($t\bar{t}Z$), which is carried out using proton-proton collision data at a center-of-mass energy of 13 TeV, recorded with the CMS detector at the LHC. This is followed by a feasibility study of W and Z boson scattering in view of the High Luminosity LHC program.

The first study is one of the most interesting analyses at the LHC as it allows to study the coupling between the top quark and Z boson. In addition, the $t\bar{t}Z$ is one of the main backgrounds for top quark pair production in association with a Higgs boson and for many beyond the standard model final states. The $t\bar{t}Z$ is measured to be $0.99^{+0.09}_{-0.08}(\text{stat.})^{+0.12}_{-0.10}(\text{sys.})$ pb, in agreement with the theoretical next-to-leading order calculation. It represents the most precise observation of this process to date attaining a 14% total relative uncertainty.

The second analysis is the scattering of W and Z bosons, which has not yet been experimentally observed. The feasibility of this process and its properties are studied with the upgraded CMS detector at the High Luminosity LHC. The results are presented in terms of the precision of the cross section, the sensitivity to the longitudinal component of the WZ scattering, and the exclusion of partial unitarization scenarios. Additionally, the performance of the upgraded detector is compared with the performance of the current detector, which will be strongly degraded due to the high radiation damage.

Samenvatting

Dit proefschrift beschrijft twee studies. De eerste studie is een meting van de werkzame doorsnede van top-quark paar productie in combinatie met een Z-boson in proton-protonbotsingen met een massamiddelpuntsenergie van 13 TeV in data genomen met de CMS detector bij LHC. Dit wordt gevolgd door een haalbaarheidsstudie van W- en Z-bosonverstrooiing bij de High Luminosity LHC.

De eerste studie is een van de meest interessante analyses bij de LHC omdat het de koppeling tussen het top-quark en het Z-deeltje mogelijk maakt. Daarnaast is het $t\bar{t}Z$-proces een van de belangrijkste achtergronden bij de productie van top-quark paren in combinatie met een Higgs-boson en bij veel eindtoestanden met deeltjes uit extensies op het standaard model. De werkzame doorsnede van de top-quark paar productie in combinatie met een Z-boson is gemeten: $0.99^{+0.09}_{-0.08}(\text{stat.})^{+0.12}_{-0.10}(\text{sys.})$ pb, in overeenstemming met de theoretische next-to-leading order berekening. Het vertegenwoordigt de observatie van dit proces bij een massamiddelpuntsenergie van 13 TeV met een precisie van 14%, de meest precieze meting tot nu toe.

De tweede analyse is de verstrooiing van W- en Z-bosonen, die nog niet experimenteel is waargenomen. De haalbaarheid van dit proces en de eigenschappen ervan worden bestudeerd met de geüpgrade CMS-detector bij de High Luminosity LHC. De resultaten worden gepresenteerd aan de hand van de precisie van de werkzame doorsnede, de gevoeligheid voor de longitudinale component van de WZ-verstrooiing, en de exclusie van gedeeltelijke unitarisatiescenario's. Daarnaast worden de prestaties van de geüpgrade detector vergeleken met de prestaties van de huidige detector, die sterk gereduceerd zal zijn vanwege de hoge stralingsschade.

Contents

List of Acroynms

ALICE	A Large Ion Collider Experiment
ATLAS	A Toroidal LHC ApparatuS
BEH	Brout-Englert-Higgs
BSM	Beyond Standard Model
CERN	European Organisation for Nuclear Research
CHS	Charged Hadron Subtraction
CKM	Cabibbo−Kobayashi−Maskawa
CMS	Compact Muon Solenoid
CP	Charge-Parity
CSC	Cathode Strip Chambers
CSV	Combined Secondary Vertex
DAQ	Data Acquisition
DT	Drift Tube
DY	Drell-Yan
ECAL	Electromagnetic Calorimeter
EWK	Electroweak
FR	Fake Rate
GSF	Gaussian Sum Filter
HCAL	Hadronic Calorimeter
HB	Hadron Barrel
HE	Hadron Endcap
HF	Hadron Forward
HL	High Luminosity
HLT	High Level Trigger
HO	Hadron Outer

IP	Impact Parameter
IVF	Inclusive Vertex Finder
JES	Jet Energy Scale
JER	Jet Energy Resolution
KF	Kalman Filter
L1	Level-1 Trigger
LEP	Large Electron-Positron Collider
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty
LO	Leading Order
MB	Muon Barrel
MC	Monte Carlo
ME	Muon Endcap
ME	Matrix Element
MVA	Multi Variate Analysis
NLO	Next-to-Leading Order
NNLO	Next-to-Next Leading Order
PDF	Parton Distribution Function
PF	Particle Flow
PS	Preshower
PU	Pile Up
RF	Radio Frequency
RPC	Resistive Plate Chamber
SC	Supercluster
SF	Scale Factor
SFOC	Same Flavour Opposite Charge
SM	Standard Model
SSB	Spontaneously Symmetry Breaking
TEC	Tracker EndCap
TIB	Tracker Inner Barrel
TID	Tracker Inner Disk
TOB	Tracker Outer Barrel
QCD	Quantum Chromodynamics
QED	Quantum Electrodynamics
QFT	Quantum Field Theory
VBS	Vector Boson Scattering

List of Figures

List of Tables

Chapter 1

Introduction

Today we know that the most fundamental constituents of the matter are elementary particles. These are briefly categorised as leptons and quarks, force carrier bosons and the Higgs boson. The theory explaining the interactions of these elementary particles is called the standard model (SM) and gained its current shape through the1960s and 1970s [1, 2]. Since then all the elementary particles of the SM, except the Higgs boson, have been discovered throughout the 20th century in different experiments.

The latest and the world's largest and most powerful particle collider, called the Large Hadron Collider (LHC), was built at the European Organisation for Nuclear Research (CERN) laboratory. The LHC project was proposed in 1984 [3] and the first collisions took place in 2010. At the time the LHC was proposed, its main purpose was to discover the missing corner stone of the SM, the Higgs boson. Furthermore precision measurements of the SM and testing the predictions of beyond the SM (BSM) theories are among the goals of the LHC. In 2012, the two general purpose experiments of the LHC reported the discovery of a boson consistent with the SM Higgs boson. This discovery completed the particle spectrum of the SM. The particle mass spectrum of the SM spans from $\sim$ 0.5 MeV to $\sim$ 173 GeV. The top quark with a mass of 172.25 GeV is the heaviest particle of the SM while the Higgs boson and the electroweak bosons, Z and $W^{\pm}$ have masses of 125.09, 91.18 and 80.38 GeV respectively. The LHC has been serving as a top quark factory due to the high instantaneous luminosity and the large top quark pair production cross section. This allows for a precise measurement of top pair production at the LHC, where its uncertainty is limited by the systematic effects. In addition, the top pair production in association with W, Z and Higgs bosons leads to one of the heaviest set of particles produced in a proton-proton collision. The measurement of these final states became possible with the increased data volumes.

This thesis presents an overview of the research I performed between 2013

and 2017 using data from the Compact Muon Solenoid (CMS) experiment at a centre of mass energy of 13 TeV at the LHC and the simulations in view of the High Luminosity LHC at a centre of mass energy of 14 TeV.

In my first year of research, I worked on a search for a heavy charged, beyond the SM, Higgs boson produced in association with a top quark, where the heavy charged Higgs decays to the SM Higgs and a $W^{\pm}$ boson. The analysis was performed using CMS proton-proton collision simulations at a centre of mass energy of 8 TeV, and the study showed that this analysis is not feasible to perform with a data volume of 20 fb^{-1}. Therefore the analysis did not advance and it is documented only as an internal note within the CMS Collaboration. The results of this study are not described in this thesis.

In my second year I got involved in the feasibility study of vector boson scattering at the High Luminosity (HL) LHC at a centre of mass energy of 14 TeV. My main contribution to this analysis was to perform the WZ scattering analysis. This study is presented in detail in Chapter 8. The results of this study is described in the technical proposal of the CMS Phase II detector (2015) and in a physics analysis summary (2016):

- Technical Proposal for the Phase-II Upgrade of the CMS Detector, Technical Report, CERN-LHCC-2015-010 [4],
- Prospects for the study of vector boson scattering in same sign WW and WZ interactions at the HL- LHC with the upgraded CMS detector. Physics Analysis Summary, CMS-PAS- SMP-14-008 [5].

The WZ scattering was studied at the centre of mass energy of 8 TeV by the CMS and ATLAS collaborations. The fiducial cross section of the WZjj process is reported with a precision of 40% by CMS [6] and an upper limit of 95% CL on the fiducial cross section was put by ATLAS [7]. At the centre of mass energy of 13 TeV, there is not yet a study of the WZ scattering and it is still being an undiscovered process at the LHC.

The main analysis of this thesis is the measurement of the cross section of top-antitop pair production in association with a Z boson ($t\bar{t}Z$) in the 3-lepton final state. I performed all the stages of this analysis which is described in detail in Chapters 6 and 7. This analysis was performed by using proton proton collisions at a centre of mass energy of 13 TeV with different data sets recorded by CMS. At first, the analysis was performed with the data sample recorded by CMS in 2015 corresponding to an integrated luminosity of 2.7 fb^{-1}. The $t\bar{t}Z$ production cross section is measured with 36% precision and with a significance of 3.6 standard deviations. The second analysis was performed with the data sample corresponding to an integrated luminosity of 12.9 fb^{-1} while the cross section is measured with a precision of 28% with 5.8 standard deviations. The final analysis performed with the available 2016 full data set that corresponds to an integrated lumi-

nosity of 35.9 fb^{-1}. The result of this analysis has a precision of of 14%. These analyses slightly differ from each other in terms of event selections.

In order to increase the sensitivity of the second and third analyses, the selection efficiencies used for the b-tagged jets and selection criteria of electrons were re-optimised. The result of this analysis have been submitted to Journal of High Energy Physics, together with the results from $t\bar{t}Z$ production in the 4-lepton final state and top-antitop pair production in association with a W boson:

- Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{13}$ TeV. The data sample corresponds to an integrated luminosity of 35.9 fb^{-1}. Submitted to JHEP, [8][1].

Previous intermediate results are published by CMS in:

- Measurement of the cross section of top quark pair production in association with a Z boson in pp collisions at 13 TeV. The data sample corresponds to an integrated luminosity of 2.7 fb^{-1}. Physics Analysis Summary, CMS-PAS-TOP-16-009 [9],
- Measurement of the top pair-production in association with a W or Z boson in pp collisions at 13 TeV. The data sample used corresponds to an integrated luminosity of 12.9 fb^{-1}. Physics Analysis Summary, CMS-PAS-TOP-16-017 [10].

This cross section was previously measured by the CMS collaboration at a centre of mass energy of 7 TeV, which was the first evidence of this process with a precision of $\sim$ 50% [11]. At 8 TeV with 20 fb^{-1} data, both the CMS and ATLAS collaborations performed the measurement of the $t\bar{t}Z$ cross section with $\sim$ 44% and $\sim$ 30% precision respectively [12], [13]. A second measurement performed on the same data set is reported by the CMS Collaboration, using multivariate techniques, and lead to the first observation of the $t\bar{t}Z$ process with $\sim$ 25% of precision [14]. These measurements were updated by the CMS collaboration at 13 TeV with different data sets and listed above, while the ATLAS Collaboration reported the measurement of the $t\bar{t}Z$ measurement with a small data set with a precision of $\sim$ 33%. The main result of this thesis, as of today, is the most precise measurement of $t\bar{t}Z$ production at 13 TeV.

The two analyses of the thesis are important studies to establish the electroweak symmetry breaking mechanism of the Higgs boson and to search

[1]The published results may slightly differ from what is presented in this thesis. The differences are minor and are results of extensive scrutiny from the CMS collaboration which are not propagated here.

for new physics in a model independent way. These two analyses have similar experimental signatures with 3 charged leptons and significant hadronic activity in the final state.

The results discussed in this thesis were presented by myself at the following conferences:

- Vector Boson Scattering prospects for High-Luminosity LHC at CMS in the WZ final state. Poster at Lepton-Photon, 17-22nd August 2015, Ljubljana [15],
- Measurement of the top quark properties in the production and decays of top anti-top pair events at CMS. Talk at Deep-Inelastic Scattering and Related Subjects, April 11-15, 2016, Hamburg[16],
- Measurement of the cross section of top quark pair production in association with a Z boson in pp collisions at 13 TeV. Talk at General Scientific Meeting 2016 of the Belgian Physical Society, May 18, Ghent,
- Vector Boson Scattering prospects for High-Luminosity LHC at CMS in the WZ final state. Poster at General Scientific Meeting 2016 of the Belgian Physical Society, May 18 2016, Ghent,
- Measurement of the cross section of ttbar production in association with a W or Z boson at 13TeV with the CMS detector. LHC Students Poster Session, February 22 2017, Geneva.

This thesis starts with a theoretical introduction to standard model of elementary particles in Chapter 2. Part I is composed of the analysis of top quark pair production in association with a Z boson: Chapter 3 gives an overview of the LHC and the CMS detector; Chapter 4 describes the simulation of the samples used which is followed by the event reconstruction methods used in CMS, in Chapter 5. Finally, Chapter 6 describes the details of the selection of the top-antitop pair production in association with a Z boson events in the 3-lepton final state, and Chapter 7 presents the results of the cross section measurement of top-antitop pair production in association with a Z boson. Part II describes the feasibility study of WZ scattering in prospect of the HL-LHC, Chapter 8 gives an introduction of the upgrade of the CMS detector for the HL-LHC followed by the study of WZ scattering. The conclusions and outlook of the analyses are given in Chapter 9.

Chapter 2

Theory

2.1 The Standard Model of Particles

The word ”physics” in Greek means nature and has its origin from a collection of manuscripts called ”The Physics” written by the philosopher Aristotle in 4th century BC. It is a collection of lectures about the most general principles of moving things and investigations of the particular contents of the universe. Our interest of the universe had started with philosophy and evolved in time with a breakthrough during the scientific revolution in 16th century where mathematical descriptive schemes led us to mechanics and astronomy. Though the history of physics is beyond the purpose of this text, our knowledge about the nature and the universe advanced from understanding the motion and dynamics of big objects like planets, stars and galaxies to a point where, today, we question what the matter is made of at the most fundamental level.

The theory explaining the interaction of the elementary particles is called Standard Model (SM). In this Chapter an overview of the main aspects of the SM is given, as a theoretical background for the work done in this thesis.

2.1.1 Elementary Particles and Their Interactions

The ”elementary particle” in physics refers to the most fundamental constituent of matter. The matter around us is made of molecules, molecules are made of atoms and atoms are made of protons, neutrons and electrons. Looking deeper to the constituents of atoms, electrons are elementary particles which in other words means it is not made of other particles, while protons and neutrons are composite particles which are made of other elementary particles.

The elementary particles can mainly be grouped in two groups according to their spin numbers: fermions and bosons. Fermions are particles which

have half-integer spin and bosons are integer spin particles.

- Elementary fermions: the building blocks of matter and they can further be categorised into leptons and quarks.
  - There are six leptons, three of them are charged and three of them are neutral. Among the charged ones the electron (e) is the well known atomic particle, while the other two are the muon (μ) and the tau (τ) that are heavier replicas of the electron. The neutral leptons are called neutrinos (ν) and come in three generations ν_e, ν_μ, ν_τ.
  - There are six quarks that are the constituents of the atomic matter. There are up (u), charm (c), top (t) quarks with electric charge of $2/3$ and down (d), strange (s), bottom (b) with electric charge $-1/3$.
- Bosons: the force carriers photon (γ),$W^{\pm}$ and Z bosons, gluons (g) and the mass-giving scalar particle, Higgs (H) boson.

The matter which surrounds us is made of electrons, up and down quarks. Protons (neutrons) are composed of uud (udd) quarks. These are called 1st generation particles as they form the everyday matter. The 2nd and 3rd generation charged leptons and quarks are called exotic matter since they are not constituents of the matter which surrounds us. They can be created in particle accelerators or produced in cosmic rays. They are not stable and decay into lighter particles, while the final decay products are the stable electrons, up and down quarks. The table of the SM particles with their mass, charge, spin informations is shown in Fig. 2.1. The table does not show the corresponding anti-particles for leptons and quarks, which have the same mass, but have opposite electric and color charges. Antiparticles are denoted with a bar above the particle symbol, i.e. $\bar{e}$ is a positron which is the anti-particle of the electron with charge $+1$.

Once the particles have been introduced, we can question how they interact with each other. The answer to this question is not as straightforward as taking two charged pith balls and observing their interaction. The motion of macroscopic objects is governed by classical mechanics, if the objects have intrinsically smaller dimensions (typically of the order of the Angstrom) their dynamics is described by quantum mechanics. If the objects gets faster, when they have speeds comparable with the speed of light, then other effects described by special relativity become visible. Finally for both small and fast things a theory that incorporates relativity and quantum principles is needed. The theory of small and fast objects are described by quantum field theory (QFT). Therefore SM describes interactions of very small and fast objects and it is based on QFT [18].

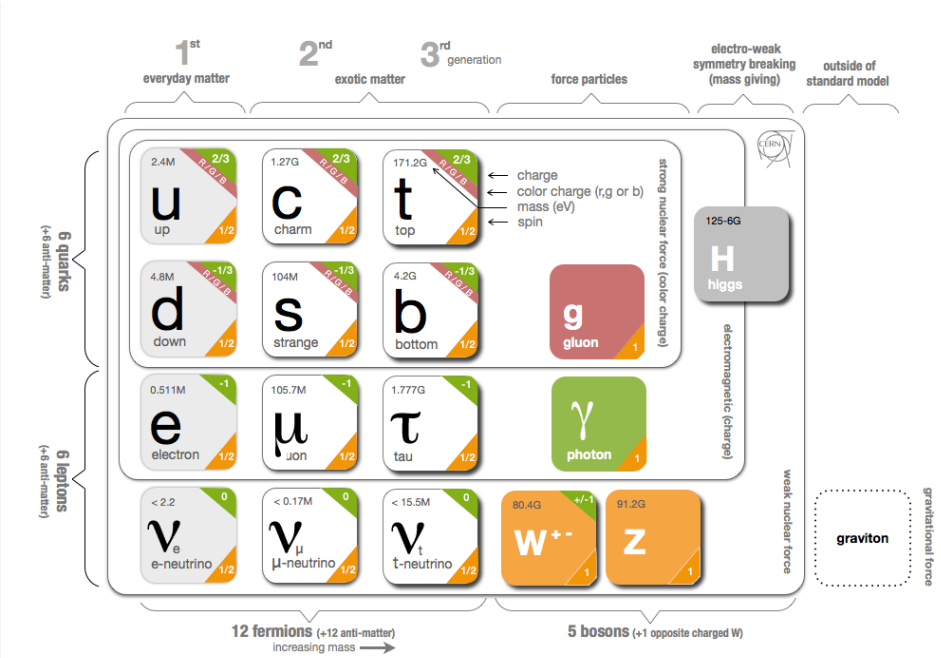

Figure 2.1: The SM particle table [17].

The elementary particle interactions are related to local symmetries[1] that means particle interactions are invariant under these symmetry transformations. In physics, the particle interactions are mostly studied by using the Lagrangian where the particles are described by fields and the interactions are introduced as a gauge symmetry of some certain Lie groups. The SM is based on the gauge symmetry of SU(3)×SU(2)×U(1) where each Lie group represents strong, weak and electromagnetic interactions respectively. The electromagnetic and weak interactions are unified as the electroweak (EWK) interaction. The generators of each symmetry group corresponds to a vector field where it is introduced in order to make the Lagrangian invariant under the gauge symmetry and called as a gauge boson.

In addition to strong, electromagnetic and weak interactions the gravitational force is what we experience everyday in our life. It holds us down to earth, keeps planets in orbit around the sun. Gravity is believed to be mediated by a force carrier particle called graviton, but this particle has not been observed so far and thus keeps being a hypothetical elementary particle. At the scale of subatomic particles the gravitational force is too weak to play a role and it is not included in SM.

[1]Local, also called gauge, symmetries are transformed by operators which are not constant in space-time.

The SM interactions and their symmetries will be introduced in the following sections.

Electromagnetic Interactions and U(1) Gauge Invariance

The electromagnetic interactions can be described by quantum electrodynamics (QED). All charged particles interact electromagnetically and the interaction is mediated by the gauge boson photon. The photon itself is neutral and does not directly interact with itself.

As a basic representation of the field theory, lets have a look to the free Lagrangian of a fermion field, denoted by ψ , with mass m and charge q and show its invariance under U(1) transformations:

$$\mathcal{L}_{free} = i\bar{\psi}\gamma_\mu\partial^\mu\psi - m\bar{\psi}\psi \tag{2.1}$$

where γ_μ are the Dirac matrices. This Lagrangian is invariant under a global transformation, where the fermion field is described by a complex field and transforms as:

$$\psi(x) \rightarrow e^{i\alpha}\psi(x) \tag{2.2}$$

where α is a real constant. While, the same does not hold true under a local U(1) transformation U(x) $= e^{i\alpha(x)Q}$, where $\alpha(x)$ is not a constant and it depends on space-time arbitrarily, and Q is the charge operator of the U(1) group (-1 for electron). The term that breaks the invariance is the derivative of the fermion field, which transforms as:

$$\partial_\mu\psi \rightarrow e^{i\alpha(x)Q}\partial_\mu\psi + ie^{i\alpha(x)Q}\psi\partial_\mu\alpha \tag{2.3}$$

In order to have an invariant Lagrangian under this transformation, the derivative is replaced by a "covariant" derivative, D_μ, that transforms covariantly like ψ itself and additionally a vector field A_μ is introduced to cancel the invariance breaking terms in the above equation.

$$D_\mu \equiv \partial_\mu - ieQA_\mu, \tag{2.4}$$

where A_μ transforms

$$A_\mu \rightarrow A_\mu + \frac{1}{e}\partial_\mu\alpha \tag{2.5}$$

Now the free Lagrangian stays invariant with the covariant derivative. This vector field, A_μ called the gauge boson, is the physical photon field. Therefore adding the kinetic energy of the photon field ($F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$), which also needs to be invariant under U(1), leads us to the QED Lagrangian:

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi + eQ\bar{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.6}$$

where the mass term $\frac{1}{2}m^2 A_\mu A^\mu$ is not allowed since it would break the gauge invariance. As a consequence, the gauge field photon is massless.

The surprising result is that by imposing a gauge invariance on the free Lagrangian, we are led to the QED Lagrangian and the massless gauge field photon [19].

The Strong Interactions and SU(3) Symmetry

All particles with color charge interact via strong interactions and are mediated by gluons. The strong interactions account for holding the proton and neutron together in an atom, as well for keeping the quarks confined in a hadron. A hadron is a composite particle made of quarks(anti-quarks). The gluon itself has a color charge, which allows for the appearance of self-interactions. None of the colored particles, quarks and gluons, can be observed as a free particle and they are confined in colorless states. This phenomena is called color confinement. Quarks come in three colors red (r), green (g) and blue (b), anti-quarks with anti-colors ($\bar{r}, \bar{g}, \bar{b}$) while gluons carry one unit of color and one unit of anti-color. The color neutral states are formed by quark-antiquark pairs (i.e. $r\bar{r}$) called mesons or groups of three quarks (anti-quarks) rgb ($\bar{r}\bar{g}\bar{b}$) called as baryons (anti-baryons). The theory of the strong interactions between quarks and gluons is called quantum chromodynamics (QCD) and described by the $SU(3)_C$ color symmetry (where C represents color). The invariance of the free Lagrangian is requested under the following SU(3) gauge transformation:

$$q(x) \rightarrow Uq(x) \equiv e^{i\alpha_k(x)T_k} q(x) \tag{2.7}$$

where q is the quark triplet denoting the three color quark states and U is an arbitrary 3×3 unitarity matrix representing the SU(3) transformation. T_k with k= 1,..8 are linearly independent traceless matrices and α_k are the group parameters. The local symmetry is restored by introducing the covariant derivative

$$D_\mu = \partial_\mu + ig_s T_k G_\mu^k \tag{2.8}$$

where g_s is the strong coupling constant, G_μ^k represents the eight gauge fields and transforms in a more complicated way compared to the photon field:

$$G_\mu^k \rightarrow G_\mu^k - \frac{1}{g}\partial_\mu \alpha_k - f_{klm}\alpha_l G_\mu^m \tag{2.9}$$

where f_{klm} are the structure constants of the group which is different from the QED case due to the non-abelian[2] structure of the SU(3) group. This leads to the self interacting gluon terms in the Lagrangian which is also

[2]The generators of the group do not commute.

	Fermions			I_3	Q	Y
Leptons	$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L$	$+1/2$ $-1/2$	0 -1	-1
	e_R	μ_R	τ_R	0	-1	-2
Quarks	$\begin{pmatrix}u\\ d\end{pmatrix}_L$	$\begin{pmatrix}c\\ s\end{pmatrix}_L$	$\begin{pmatrix}t\\ b\end{pmatrix}_L$	$+1/2$ $-1/2$	$2/3$ $-1/3$	$1/3$
	u_R	c_R	t_R	0	$+2/3$	$4/3$
	d_R	s_R	b_R		$-1/3$	$-2/3$

Table 2.1: Overview over the elementary fermions of the SM and their quantum numbers: I_3 third component of the weak isospin, Q electric charge and Y weak hypercharge.

different than the photon field. Adding the gauge invariant kinetic term for each of the gluon fields, the gauge invariant QCD Lagrangian is the following:

$$\mathcal{L}_{QCD} = \bar{q}(i\gamma^\mu \partial_\mu - m)q - g(\bar{q}\gamma^\mu T_a q)G^a_\mu - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a. \tag{2.10}$$

As in the case for U(1) gauge invariance, acquiring the Lagrangian to be invariant under color gauge transformations leads us to 8 self interacting massless gluon fields.

The Weak Interactions and Electroweak Unification

The weak force accounts for the known nuclear beta decay, muon decay and it is mediated by the massive $W^\pm$ and Z bosons. Unlike the electromagnetic and strong interactions, only left handed fermions and right handed anti-fermions interact weakly. This breaks the chiral symmetry and as a result the massive gauge fields, unlike the massless photon and gluon fields, break the gauge invariance. As a consequence, the gauge symmetry of the weak interactions is more complicated compared to the U(1) and SU(3) symmetries.

In order to describe the weak interactions of fermions, the electromagnetic and weak interactions are unified as electroweak interactions [20]. The electroweak interaction is invariant under the $SU(2)_L \times U(1)_Y$ weak isospin and hypercharge symmetry, L stands for $left$ meaning acting only on the left handed fermions and Y is the weak hypercharge defined as $Q = I_3 + \frac{Y}{2}$, I_3 is the third component of weak isospin [3]. Left handed fermions are grouped into doublets of weak isospin $I_3 = \pm 1/2$ and right handed fermions are

[3] Isospin is introduced as a global-flavor symmetry in strong interactions relating up and down quarks.

isospin singlets with $I_3 = 0$. The overview of the representations of the fermions and their quantum numbers are shown in Tab. 2.1.1. In $SU(2)_L \times U(1)_Y$ the left handed and right handed fermions transform different:

$$\begin{aligned} L &\rightarrow e^{i\alpha_k(x)T_k + i\beta(x)\frac{Y}{2}} L \\ R &\rightarrow e^{i\beta(x)\frac{Y}{2}} R \end{aligned} \tag{2.11}$$

where $T_k = \tau_k/2$, $(k = 1, 2, 3)$ is the generators of weak isospin group $SU(2)_L$, τ_i are the Pauli-matrices, $Y/2$ is the generator of the hypercharge group U(1) and R represents the right handed fermions.

As in U(1) and SU(3) representations, we can introduce the vector fields to ensure the gauge invariance: W^k_μ, $k = 1, 2, 3$ is introduced for the $SU(2)_L$ and a single vector field B_μ for the $U(1)_Y$. Then the covariant derivative is:

$$D_\mu = \partial_\mu + igT_k W^k_\mu + i\frac{g'}{2} Y B_\mu \tag{2.12}$$

with couplings g and g' for the $SU(2)_L$ and $U(1)_Y$ respectively.

The $SU(2)_L$ is a non-abelian group, as SU(3), that the W_k vector fields transform similar to the gauge bosons as in Eq. 2.9 and the B vector field transform as in Eq. 2.5. Adding the kinetic terms of the gauge bosons to the free Lagrangian leads us to the electroweak Lagrangian:

$$\mathcal{L}_{EWK} = i\bar{L}\gamma^\mu D_\mu L + i\bar{\psi_R}\gamma^\mu D_\mu \psi_R - \frac{1}{4} W^k_{\mu\nu} W^{k,\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \tag{2.13}$$

The electroweak Lagrangian is invariant with massless vector bosons, however it is known that the $W^\pm$ and Z bosons are massive. The masses of the vector bosons need to be added in the theory ensuring the gauge invariance. This happens by the electroweak symmetry breaking mechanism, so called the Brout-Englert-Higgs (BEH) mechanism [1, 2].

2.1.2 Spontaneous Symmetry Breaking and the BEH Mechanism

Up to now, we have shown that the mass terms of the gauge bosons are not allowed in a gauge invariant theory. As a consequence massive gauge bosons will break the symmetry. In order to allow massive gauge bosons while keeping the Lagrangian invariant under the presented gauge symmetries, we need to introduce the spontaneously symmetry breaking (SSB) phenomenon. The SSB is achieved by adding a scalar field to the Lagrangian, for which the non-zero vacuum expectation values (ground state) break the symmetry. The choice of the field for an SU(2) gauge symmetry is a doublet of complex scalar fields:

$$\Phi = \begin{pmatrix} \varphi^\dagger \\ \varphi \end{pmatrix} = \frac{1}{\sqrt{(2)}} \begin{pmatrix} \varphi^1 + i\varphi^2 \\ \varphi^3 + i\varphi^4 \end{pmatrix} \tag{2.14}$$

and the SU(2) invariant Lagrangian with the simplest renormalizable potential is:

$$V(\Phi^\dagger\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2,$$
$$\mathcal{L} = (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\Phi^\dagger\Phi) \tag{2.15}$$

where the covariant derivative is defined in Eq.2.12.

There are two possible forms of this potential depending on the sign of μ^2. If $\mu^2 > 0$, the minimum of the potential can be set at $< \Phi >= 0$. This represents a system of four scalar particles (φ_i of Eq.2.14) each with a mass μ interacting with 3 massless gauge bosons W^k_μ and it does not break the symmetry. This is illustrated in the top plot in Fig. 2.2. We are interested in the case $\mu^2 < 0$. The ground state minima is given by

$$\Phi^\dagger\Phi = \frac{1}{2}(\varphi_1^2 + \varphi_2^2 + \varphi_3^2 + \varphi_4^2) = -\frac{\mu^2}{2\lambda}. \tag{2.16}$$

The corresponding potential is shown in the bottom plot in Fig. 2.2 for one field. At $< \Phi >= 0$ the ground state is not stable, illustrated in the bottom left plot in Fig. 2.2, and the field will move from the origin point (from higher potential energy) to the actual minimum state. Choosing one of the ground state breaks the symmetry, as illustrated in the bottom right plot in Fig. 2.2 . This ground state is also called the vacuum expectation value (v) and equal to $\pm\sqrt{\frac{\mu^2}{\lambda}}$. The Lagrangian symmetry is broken by the choice of one of the ground states, it is either $+v$ or $-v$, where the Lagrangian is not symmetric. The field needs to conserve the U(1) symmetry and breaks $SU(2)_L$. Therefore the field can be fixed to a minimum energy position by choosing $\varphi_{1,2,4} = 0$ and $\varphi_3^2 = \frac{-\mu^2}{\lambda} \equiv v^2$ and can be parametrised by h(x) which represents the fluctuations of this minimum:

$$\Phi = e^{\tau_i\theta^i(x)/v}\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v + h(x)\end{pmatrix}. \tag{2.17}$$

Here h(x) is the BEH field, $\tau_{1,2,3}$ are the generators of $SU(2)_L$ and $\theta_{1,2,3}$ are the massless Goldstone bosons. According to the Goldstone theorem, the spontaneously broken symmetry leads to massless scalars as many as the broken generators. The $SU(2)_L$ symmetry allows to rotate away any dependence on $\theta_i(x)$. Choosing the unitarity gauge $\theta_i(x) = 0$, eliminates the θ_i fields in the Lagrangian, that Goldstone bosons are eaten by the three gauge bosons that require masses and give the longitudinal components[4] to the massive gauge bosons.

[4]Massless bosons have two transverse polarisation states$(-,+)$, while massive bosons can have the third polarisation which is the longitudinal component

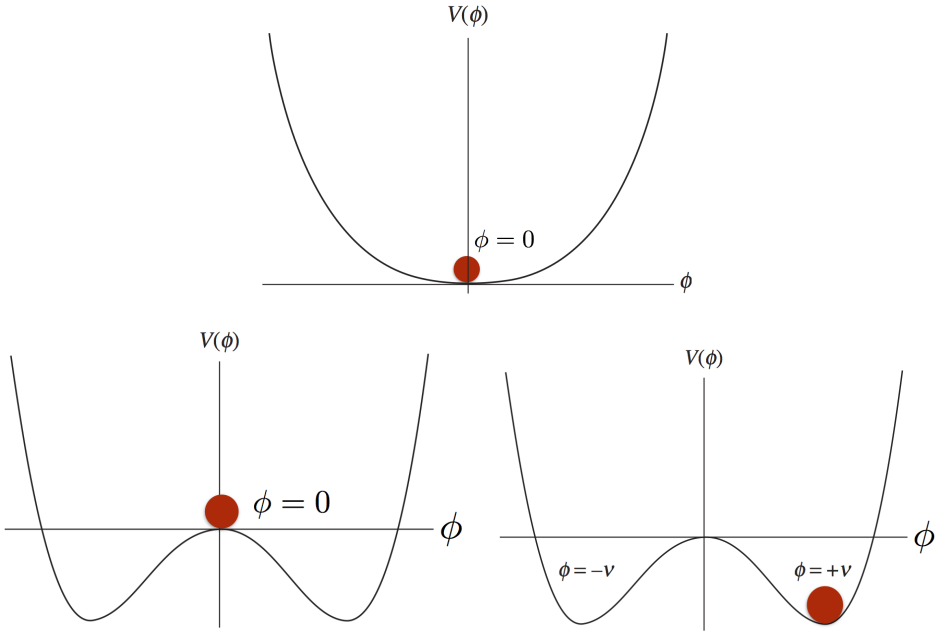

Figure 2.2: The potential of the Lagrangian: top without symmetry breaking case for $\mu^2 > 0$; bottom symmetry breaking for case for $\mu^2 < 0$. The bottom left shows the minimum of the potential is not at $< \Phi >= 0$, the bottom right shows the minimum of the potential which breaks the symmetry

The BEH potential of the SSB Lagrangian takes the following form:

$$V(\Phi^\dagger\Phi) = \frac{1}{2}(2\lambda v^2)h(x)^2 + \lambda v h(x)^3 + \frac{\lambda}{4}h(x)^4 - \frac{\lambda}{4}v^4. \tag{2.18}$$

The BEH potential has quadratic, cubic and quartic terms of the BEH field. The first term is the mass term of the BEH field:

$$m_H = \sqrt{2\lambda}v = \sqrt{2}|\mu| \tag{2.19}$$

and it depends on the self BEH coupling λ and the v. The cubic and quartic terms correspond to self interactions of the BEH field, and the last term is a constant.

The other part of the Lagrangian of Eq.2.15 is the $|D_\mu\Phi|^2$ term. When we insert the new scalar field with the covariant derivative, it becomes:

$$(D^\mu\varphi)^\dagger(D_\mu\varphi) = \frac{1}{2}|\partial_\mu h(x)|^2 + \frac{1}{8}v^2\left[g^2(W_1^2 + W_2^2) + (gW_3 - g'B_\mu)^2\right] + \mathcal{O}(h(x)). \tag{2.20}$$

Here the first term is the kinetic term of the BEH field while the last term has the interactions of the BEH field with the gauge boson. We will focus on

the second term in the Lagrangian which gives the mass terms of the gauge bosons. We can rewrite this term of the Lagrangian in terms of the known $W^{\pm}, Z$ and A bosons as:

$$\frac{1}{8}v^2[g^2(W^+)^2 + g^2(W^-)^2 + (g^2 + g'^2)Z_\mu^2 + 0 \cdot A_\mu^2], \tag{2.21}$$

where

$$W^{\pm} = \tfrac{1}{\sqrt{2}}(W_1 \pm iW_2) \text{ with } M_{W^{\pm}} = \tfrac{1}{2}vg$$

$$Z_\mu = \tfrac{1}{\sqrt{g^2+g'^2}}(gW_3 - g'B_\mu) \text{ with } M_Z = \tfrac{1}{2}v\sqrt{(g^2 + g'^2)} \tag{2.22}$$

$$A_\mu = \tfrac{1}{\sqrt{g^2+g'^2}}(g'W_3 + gB_\mu) \text{ with } M_A = 0.$$

The ratio of the M_W and M_Z bosons $g/\sqrt{g^2 + g'^2}$ is equal to the cosine of the weak mixing angle, $cos(\theta_W)$ [21]. The weak mixing angle is a parameter of the SM that rotates the W_3, B_μ vector boson plane producing the Z and A_μ bosons by SSB. Additionally it relates the couplings as:

$$e = gsin\theta_W = g'cos\theta_W. \tag{2.23}$$

The experimental measurements of the M_W, M_Z and θ_W confirm the above relations.

Up to now, using the gauge invariance of the theory, we showed how the W and Z bosons gain their mass while the photon remains massless with the addition of the BEH field. But we still need to discuss how fermions acquire their mass. In Section 2.1.1 we have shown how the fermion fields transform under $SU(2)_L \times U(1)_Y$ rotations. The mass terms of the fermions are not allowed since left handed fermions form an isospin doublet and right handed fermions form isospin singlets and terms like $m[\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L]$ are not gauge invariant. Therefore a singlet term of $SU(2)_L \times U(1)_Y$ is needed for an invariant Lagrangian mass term. This can be done by using the BEH doublet introduced in this section:

$$\mathcal{L}_{fermions} = \lambda_f[\bar{\psi}_L\varphi\psi_R + \bar{\psi}_R\varphi\psi_L] \tag{2.24}$$

for electron this term becomes

$$-\frac{\lambda_e(v+h)}{\sqrt{2}}[\bar{e_L}e_R + \bar{e_R}e_L] = -\frac{\lambda_e v}{\sqrt{2}}\bar{e}e - \frac{\lambda_e}{\sqrt{2}}h\bar{e}e \tag{2.25}$$

where e_L, e_R refer to the left and right handed electrons. The first term gives the mass term of the electron, $\lambda_e v/\sqrt{2}$ and the second term describes the interactions between the BEH field and the fermions. The λ parameter is described as the Yukawa coupling and expressed as $\lambda_f = \sqrt{2}(\frac{m_f}{v})$ and the

coupling of the fermion to the BEH field is $\lambda_f/\sqrt{2} = m_f/v$ where $v \sim 246$ GeV, which is proportional to the mass of the fermion. This mass term only gives mass to 'down' type of leptons, while keeps the 'up' type of leptons, neutrinos, massless[5]. In order to have mass terms for the up type quarks, an additional term is needed in the Lagrangian. This is done by introducing the charge conjugate representation of the BEH doublet, which under SU(2) rotations transforms as the original BEH field:

$$\tilde{\varphi^C} = -i\tau_2\varphi^\star = \sqrt{\frac{1}{2}}\begin{pmatrix} v+h \\ 0 \end{pmatrix}, \tag{2.26}$$

where $\tilde{\varphi^C}$ is the charge conjugate representation of the BEH doublet. The mass term of the up-type fermion becomes:

$$\mathcal{L}_{up} = \lambda_q[\bar{U_L}\tilde{\varphi^C}U_R + \bar{U_R}\tilde{\varphi^C}U_L], \tag{2.27}$$

where U represents the up type fermions. This mass term has the same form as the down fermions with the corresponding Yukawa couplings. All the mass terms of the SM particles can be expressed in terms of the vacuum expectation value v and the coupling constants: g, g', λ_i where the Yukawa couplings, λ_i are different for each lepton and quark and zero for neutrinos in the SM. The measurements of the Yukawa couplings to fermions and weak vector bosons are shown in Fig. 2.3 as a function of the particle mass.

Finally we can gather all the ingredients of the SM, SU(3)xSU(2)$_L$ $xU(1)_Y$ and summarize all the interaction and mass terms in Lagrangian:

$$\begin{aligned}
\mathcal{L} = & \underbrace{-\frac{1}{4}W_{\mu\nu}\cdot W^{\mu\nu} - \frac{1}{4}B_{\mu\nu}\cdot B^{\mu\nu} - \frac{1}{4}G_{\mu\nu}\cdot G^{\mu\nu}}_{W^\pm,Z,\gamma \text{ and gluon kinetic energies and self interactions}} \\
& + \underbrace{\ell_L\gamma^\mu(i\partial_\mu - g\frac{1}{2}\tau^i W^i_\mu - g'\frac{Y}{2}B_\mu)\ell_L + q_L\gamma^\mu(i\partial_\mu - g\frac{1}{2}\tau^i W^i_\mu - g'\frac{Y}{2}B_\mu - g_s T^k G^k_\mu)q_L}_{\text{left handed fermion kinetic energies and their interactions with } W^\pm,Z,\gamma \text{ and gluons}} \\
& + \underbrace{\ell_R\gamma^\mu(i\partial_\mu - g'\frac{Y}{2}B_\mu)\ell_R + q_R\gamma^\mu(i\partial_\mu - g'\frac{Y}{2}B_\mu - g_s T^k G^k_\mu)q_R}_{\text{right handed fermion kinetic energies and their interactions with } \gamma \text{ and gluons}} \\
& + \underbrace{|(i\partial_\mu - g\frac{1}{2}\tau^i W^i_\mu - g'\frac{Y}{2}B_\mu)\varphi|^2 - V(\varphi^\dagger\varphi)}_{W^\pm,Z,\gamma \text{ and BEH masses and coupling}} - \underbrace{(\lambda_f\bar{\ell_L}\varphi\ell_R + \lambda_g\bar{q_L}\tilde{\varphi^C}q_R + \text{h.c.})}_{\text{fermion masses and couplings to BEH}}
\end{aligned} \tag{2.28}$$

where ℓ is used for leptons and q for quarks.

[5]Despite experimental evidence for neutrino oscillations, which implies non-zero neutrino masses, the SM does not predict non-zero masses for neutrinos in a natural way.

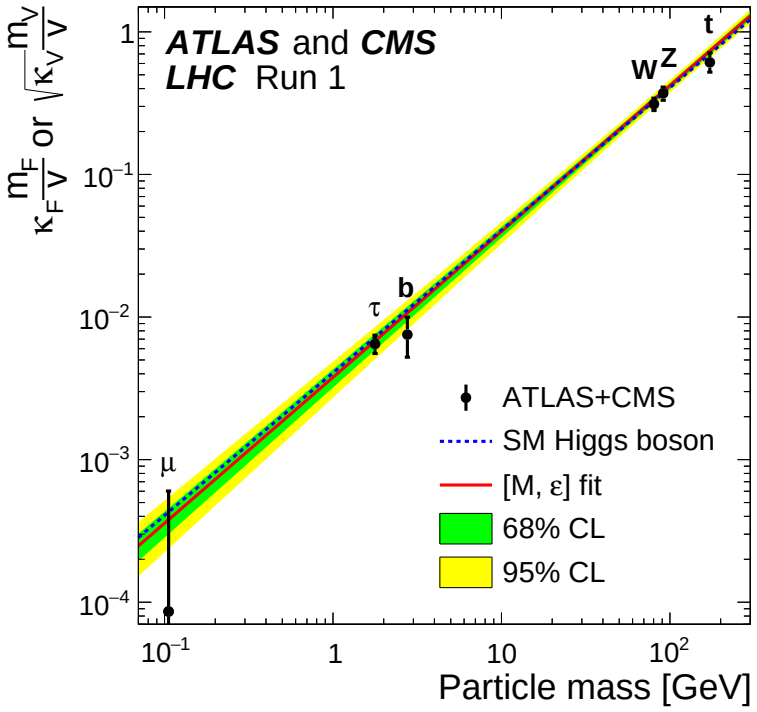

Figure 2.3: Best fit values of the Yukawa coupling as a function of particle mass for the combination of ATLAS and CMS data at $\sqrt{7}$ and $\sqrt{8}$ TeV. m_F, m_V corresponds to fermion and vector boson masses and $v = 246$ GeV is the vacuum expectation value of the BEH field. k_F, k_V are the coupling modifiers and expected to be equal to unity in SM. The dashed (blue) line indicates the predicted dependence on the particle mass in the case of the SM Higgs boson. The solid (red) line indicates the best fit result with the corresponding 68% and 95% CL bands, [22].

2.2 Beyond the Standard Model

Over the past decades the predictions of SM were heavily tested in various experiments. So far, no direct experimental evidence of deviation from SM predictions was observed. An example where one can appreciate the experimental data described very well with the SM predictions is given in Figure 2.4. The figure presents a summary of measured production cross sections for various processes at 7, 8, 13 TeV of collision energies compared to theoretical predictions. Despite the great success of SM, there exist a number of puzzling phenomena of nature which cannot be shed light within the framework of SM.

One of the puzzles come from neutrino physics. The experimental observation of the neutrinos [23],[24] imply at least two of the neutrinos to have non-zero mass, while in its most natural form SM treats neutrinos massless. The neutrino changes flavour as it propagates and this is possible only if

there is a mass difference between the three flavours of neutrinos. The exact masses of the neutrinos are yet unknown.

From cosmology it is known that only about 5% of the universe is composed of quarks and leptons which can be explained by SM. The $\sim$ 25% of the universe is composed of dark matter which does not interact with the baryonic matter via electromagnetic interactions that it is invisible for detection. It is known to exist due to its gravitational interaction and its effects on galactic rotations of stars in galaxies [25]. There are many theoretical models and experimental searches to explain dark matter and its origin, for a detailed review about the searches in LHC see [26]. In addition, the dark energy, assuming the current cosmological models are correct, contributes approximately 70% of the universe. It is an unknown form of energy which is hypothesised to explain the accelerated expansion of the universe [27].

Another puzzle awaiting to be understood in particle physics is the matter and anti-matter asymmetry of the universe. According to the Big Bang theory, equal amount of matter and anti-matter was produced in the early universe. However the universe is dominated by matter. The necessary conditions that cause this asymmetry, proposed by Sakharov [28], are baryon number violation (baryogenesis), charge-parity (CP) violation and interactions outside of thermal equilibrium that the previous conditions can happen. CP violation is allowed by SM and from experiments it is known to happen in weak interactions, however its rate is not large enough to explain the matter asymmetry of the universe[29].

Besides the puzzles due to observations and experiments, some features of the theoretical structure of SM lack understanding. The theory does not predict the particle masses and also does not explain the hierarchy among their masses or why there are three generations of fermions. The SM works at the electroweak scale (order of 100 GeV) and the gravity is negligible at these energies. At Planck scale, order of 10^{19} GeV, the gravity effects will become important. But SM can not include gravity that at this scale it will no longer be valid, which is the cutoff scale of the SM. In order to have SM to be valid up to this scale, the radiative corrections are needed to be applied to the Higgs boson mass. These corrections are quadratic in cutoff scale and they need to cancel out to have Higgs boson mass at the EWK scale. This is assumed to happen unnaturally through fine tuned SM parameters that SM does not require the parameters to be finely tuned. This is called as the hierarchy problem. This problem is solved in various new physics models, such as Supersymmetry with introducing partners of the SM particles in the theory which somewhat naturally cancels out large radiative corrections on the Higgs boson mass saving it from divergences.

All these hint towards the fact that the SM may not the be ultimate theory that it is an effective theory at low energies while at higher energies a more

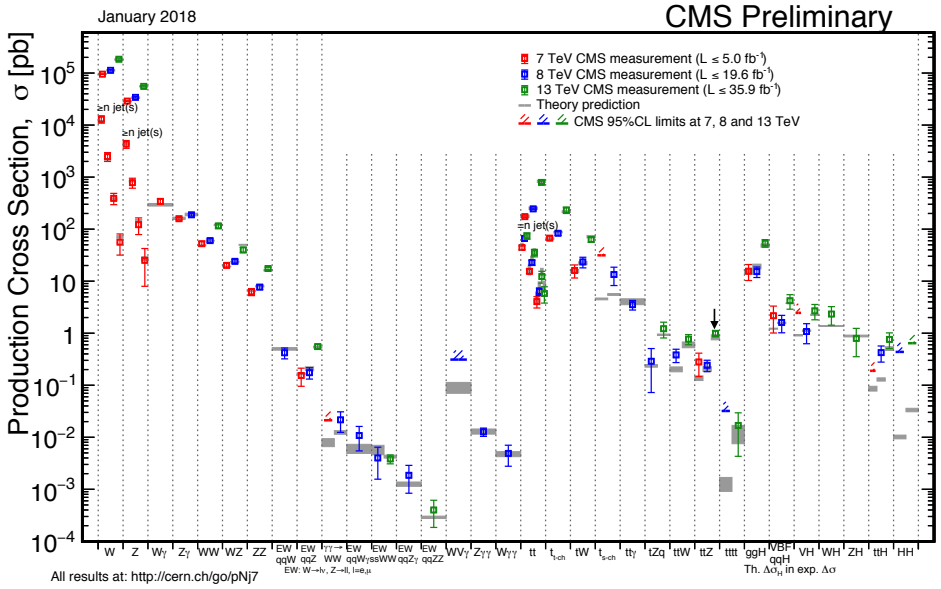

Figure 2.4: Summary of the cross section measurements of Standard Model processes with the CMS experiment with 7, 8 and 13 TeV data.

fundamental theory should be governing the nature. Thus particle physicists are in continuous hunt for evidence of such a new theory. In this thesis the two subjects investigated serve this purpose in an indirect way.

2.3 Top Quark

The top quark was discovered by the CDF and D0 experiments at the $p\bar{p}$ collider at Tevatron in 1995 [30, 31]. Since its discovery, it has been one of the main study in particle physics due to its properties. It is by far the heaviest of all known quarks and of all discovered elementary particles. Its mass is $m_t = 173.1 \pm 0.6$ GeV,[6] [32] and its Yukawa coupling to the BEH boson is order of unity. The large value of the top quark Yukawa coupling leads to the question whether the top quark has a significant role in electroweak symmetry breaking. In electroweak theory, the top quark appears in higher order loop diagrams making m_t as crucial parameter in the theory [33]. In addition to this, in BSM theories, such as top-color model, this quark plays a special role in the electroweak SSB sector, for a complete description please see [34]. Additionally, in several BSMs new particles decay to top quarks or new particles can be produced from its decays.

As it is much massive than the W boson, it is the only quark that can decay into a real W boson and a b quark, that gives the opportunity to measure

[6] Throughout this thesis, the natural units are used, such that $\hbar = c = 1$.

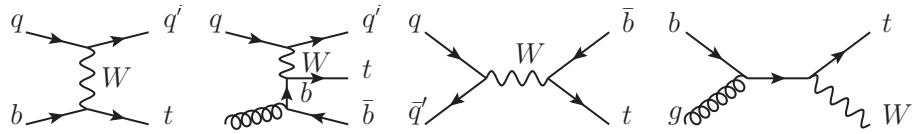

Figure 2.5: Lowest order Feynman diagrams for single top quark production. The first two left: $t-$channel production, the last two right; $s-$channel production[33]

the top quark to W boson coupling. Moreover its lifetime is shorter than the hadronization time, meaning that top quark decays before it can hadronize unlike the other quarks. This allows to study directly the bare quark properties. Therefore, since its discovery at the Tevatron, top quark properties and its interactions are being studied in detail. For a full review please see [33].

2.3.1 Top Quark Production and Decay

Top quark can be produced alone or in top anti-top quark pairs. The single top quarks are produced in electroweak interactions with Wtb vertex in two different modes, shown in Fig. 2.5:

- t channel mode:[7] a W boson is scattered of by a b quark coming from the constituents of the initial proton or coming from a gluon splitting and becomes a top quark.
- s channel mode:[8] one case is W boson is produced via the interaction of a quark and an anti-quark, i.e. $u\bar{d}$ and decays into $t\bar{b}$ and the other is top quark produced in association with a W boson via the interaction of an initial gluon with a b quark.

The single-top production rate measurements are important since the single top production is sensitive to the Wtb vertex in different ways and also allows the direct measurement of the Cabibbo$-$Kobayashi$-$Maskawa matrix (CKM) which is expected to be almost 1 in SM. Therefore, deviations in the measurements of the CKM matrix and the Wt couplings would indicate new physics. Another importance of the single top production is that it is produced left-handed in its rest frame. Since top quark decays before it hadronizes, its polarisation state is transferred to its decay products. An observable variable sensitive to the top quark polarisation is the forward-backward asymmetry. It is expressed in terms of the angle between the lepton, which is the decay product of the W coming from the top quark,

[7] It is one of the Mandelstam variables, also called time channel, which is a Lorentz invariant variable used in Feynman calculus. If 1 and 2 are the incoming and 3 and 4 are the outgoing particles t variable is defined as: $(p_1 - p_3)^2 = (p_4 - p_2)^2$

[8] It is one of the Mandelstam variables, also called space channel. As in the t variable definitions, s variable is defined as: $(p_1 + p_2)^2 = (p_3 + p_4)^2$. The s variable is commonly used in particle physics to define the center of mass of the system, i.e. LHC is running at $\sqrt{s}$.

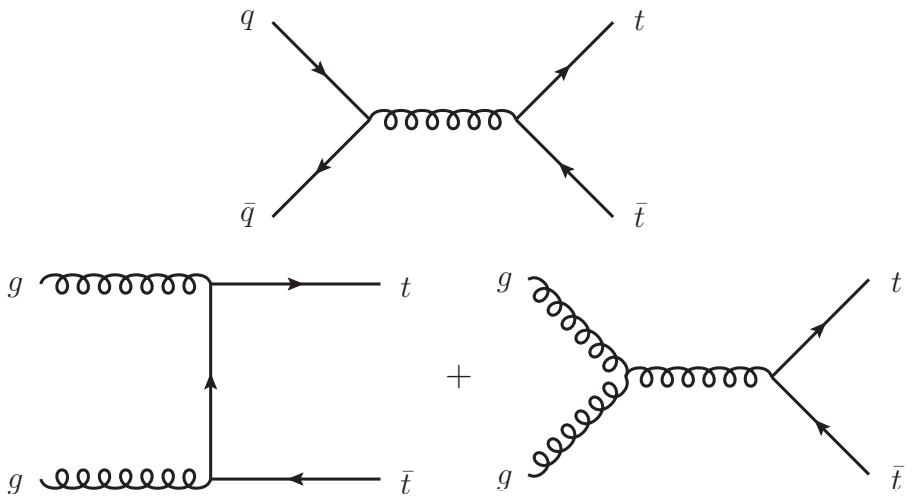

Figure 2.6: Lowest order Feynman diagrams for top pair production. The top plot shows the quark anti-quark pair while the bottom ones gluon-gluon fusion.

and the light quark, which is the other outgoing quark in single top quark production in t channel (denoted by q' in the first two left plots in Fig.2.5). The theoretical predictions of this asymmetry with the measured ones are in agreement while deviations could be a sign of BSM. Yet the measurements do not show any deviations from the SM [35, 36].

At hadron colliders, the top anti-top pairs are predominantly produced via strong interactions. At $p\bar{p}$ colliders, like at Tevatron, the top pairs are dominantly produced in quark anti-quark annihilation while in pp collisons, like at LHC, the dominant production is through gluon-gluon fusion. The corresponding leading order Feynman diagrams are shown in Fig. 2.6

The decays of top quark pairs can be classified according to the decay mode of the W boson, since top quark always decays to a W boson and a b quark. This is constrained by the CKM mechanism, where top quark decaying to other generations is strongly suppressed in SM and this is confirmed in all measurements up to now.

- Hadronic channel: $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'bq\bar{q}'\bar{b}$ both W bosons decay into a quark anti-quark pair. The branching fraction of this channel is $\sim$ 46%. The multiple final state jets makes it difficult to reconstruct the two hadronically decaying W bosons. Therefore, this channel is the experimentally the most challenging to study.
- Semi-leptonic channel: $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'b\ell\nu b$ or $\bar{\ell}\bar{\nu}bq\bar{q}'\bar{b}$ one W boson decays leptonically and the other one into a quark anti-quark pair. The branching fraction of this channel is $\sim$ 44% considering electron, muon and tau final states.

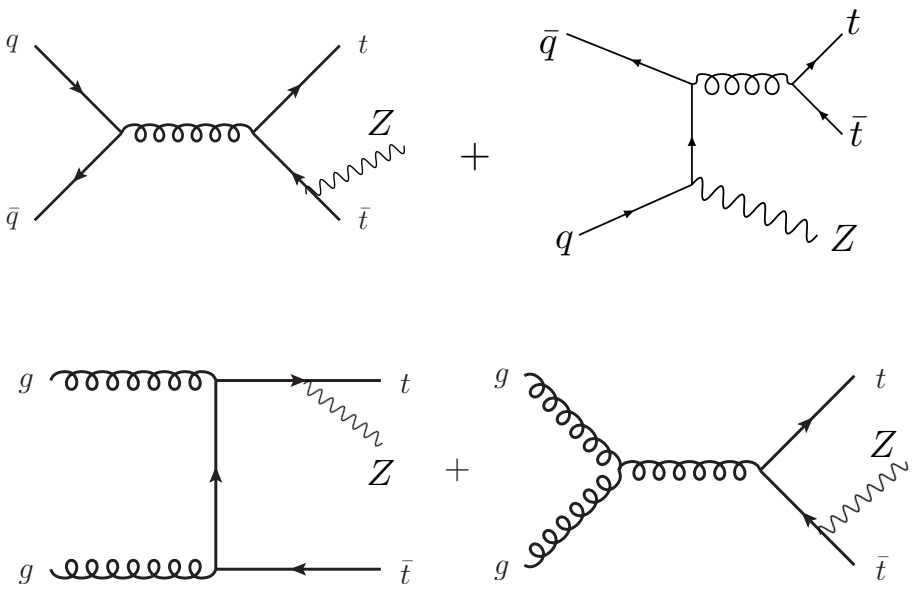

Figure 2.7: Lowest order Feynman diagrams for the tt̄Z production at hadron colliders where a Z boson is radiated from a top quark or or from the initial quarks.

- Di-lepton channel: $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \bar{\ell}\nu b\ell\bar{\nu}\bar{b}$ both W bosons decay into a lepton and a neutrino(leptonically). The branching fraction of this channel is $\sim 10\%$ considering electron, muon and tau final states. This final state is experimentally the easiest and cleanest final state with the two energetic opposite sign leptons in the final state. The two neutrinos escape detection that appears as large missing transverse energy.

2.3.2 Top Quark Pair Production in Association with Electroweak and Higgs Bosons

Top quark pair production in association with $W^{\pm}$,Z and H bosons is one of the interesting final states in experiments due to several reasons. The top quark coupling to the Higgs boson can be studied in tt̄H production. A direct measurement of this coupling is important for the confirmation of the SM and any deviations from the theory may be sign of new physics.

The Z boson radiated off one of the top quarks is sensitive to the t-Z coupling. The representative lowest order diagrams are shown in Fig. 2.7. The decay channels of tt̄Z production can be categorised in terms of the number of leptons in the final state. Referring to the top pair decay channels described in the previous section, we can add up the Z boson decays together with tt̄:

- 0-lepton: tt̄ and Z boson both decay hadronically. This final state has

the highest branching ratio $\sim$ 50%, while experimentally it has a very difficult final state due to the multi-jet backgrounds that make the reconstruction of the final state objects among many jets very challenging.

- 1-lepton: $t\bar{t}$ decays semi-leptonically and Z boson decays hadronically. This final state has a high branching ratio $\sim$ 35%, while experimentally it has a very difficult final state as in the case of 0 lepton final state. These two channels have not been used in any analysis yet due to their complicated final states.

- 2-lepton: $t\bar{t}$ decays hadronically and Z boson decays to opposite sign leptons. This final state has a branching ratio $\sim$ 10%, as in hadronic $t\bar{t}$ case this channel is not favoured due to multi-jet background final states.

- 3-lepton: $t\bar{t}$ decays semi-leptonically and Z boson decays to opposite sign leptons. This final state has a very low branching ratio, $\sim$ 3%. But thanks to its topology with 3 leptons in the final state and to the higher cross section with respect to its background, this is the most promising channel to obtain a high *signal/background* ratio.

- 4-lepton: both $t\bar{t}$ and Z boson decay leptonically. Even though this final state has the lowest branching ratio $<$ 1%, this final state has a very clean experimental signature with 4 leptons.

The $t\bar{t}Z$ production has a very similar topology to $t\bar{t}H$, which makes $t\bar{t}Z$ as one of the main backgrounds of $t\bar{t}H$ in 2 and 3 lepton final states. It is also one of the main backgrounds for several BSM searches with multi-leptons. Additionally, $t\bar{t}Z$ allows the direct measurement of the t-Z coupling which makes it crucial to understand and measure this final state precisely.

Furthermore, $t\bar{t}Z$ and $t\bar{t}H$ measurements can be interpreted in the context of an effective field theory to constrain the Wilson coefficients [37]. In this approach, cross section measurements can be used to search for new physics in a model independent way where the SM Lagrangian is extended through higher order operators [38, 39]. In order to study the effects of new physics, the expected cross sections are calculated as a function of the Wilson coefficients which parametrise the strength of the new physics interactions. Some of these operators affect $t\bar{t}Z$, $t\bar{t}W$ and $t\bar{t}H$ processes separately and simultaneously. The details of these operators and Wilson coefficients have been reported in [40, 41]. The most recent experimental constraints from $t\bar{t}W$ and $t\bar{t}Z$ measurements are reported in [14, 42]. The ttZ production in 3-lepton final state is studied in this thesis by using the LHC data of pp collisions at a center of mass energy of 13 TeV. This will be described in Chapter 6-7. The result of this measurement is denoted with the black arrow in Fig. 2.4.

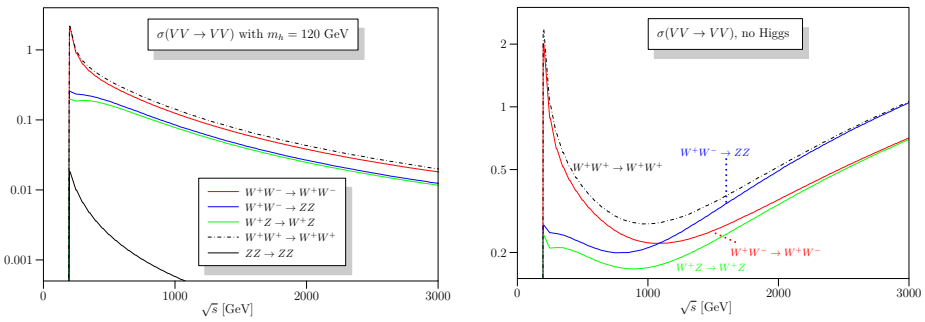

Figure 2.8: Cross sections (in nanobarns) for the scattering processes of longitudinal weak gauge bosons: SM with a 120 GeV (left) and SM without the BEH (right) boson, taken from [43]

2.4 Vector Boson Scattering

The study of Vector Boson Scattering (VBS) plays an important role in the EWK sector of the SM. As it was explained in Sec. 2.1.2, $W^{\pm}, Z$ bosons become massive by the EWK symmetry breaking mechanism. The vector bosons acquire mass through their coupling to the BEH field and gain their longitudinal polarisation components. The transverse components of the vector bosons correspond to the vector bosons that were present in the theory before the SSB, while the longitudinal components arise for massless bosons which absorb the Goldstone bosons ($\theta_{1,2,3}$ in Eq. 2.17) in SSB. According to the Goldstone boson equivalence theorem [44, 45, 46], the longitudinal components of the vector boson scattering cross section decrease linearly with the scattering energy, $\sigma \sim 1/s$, in high energy limit $s \gg m_H$. This finite cross section value is conserved due to the cancellations of the Higgs and the Goldstone diagrams. In a theory with no Higgs boson or with different Higgs vector boson couplings values than the SM ones, the cancellations are not complete and the cross section diverges. This is called as the unitary violation and happens at an energy around 1.2 TeV [47, 48]. The longitudinal VBS cross sections, as a function of the energy, in SM and in the case with no Higgs boson is illustrated in Fig. 2.8. In this plot, on the left, it can be seen that the cross section decreases with increasing energy (the SM case) while on the right plot the cross section increases after $\sim$ 1.2 TeV (the SM with no Higgs boson).

The discovery of the Higgs boson at CERN [49, 50] has so far shown SM-like properties. Thus, this boson may be responsible for preserving the unitarity at high energies. Therefore the confirmation of this through VBS studies provides a model-independent test of the SM. Moreover, several BSM theories modify the vector boson couplings that their effects can be parametrised in an effective field theory approach [51]. The VBS topology allows tests for the presence of anomalous quartic gauge couplings [52] via its triple and

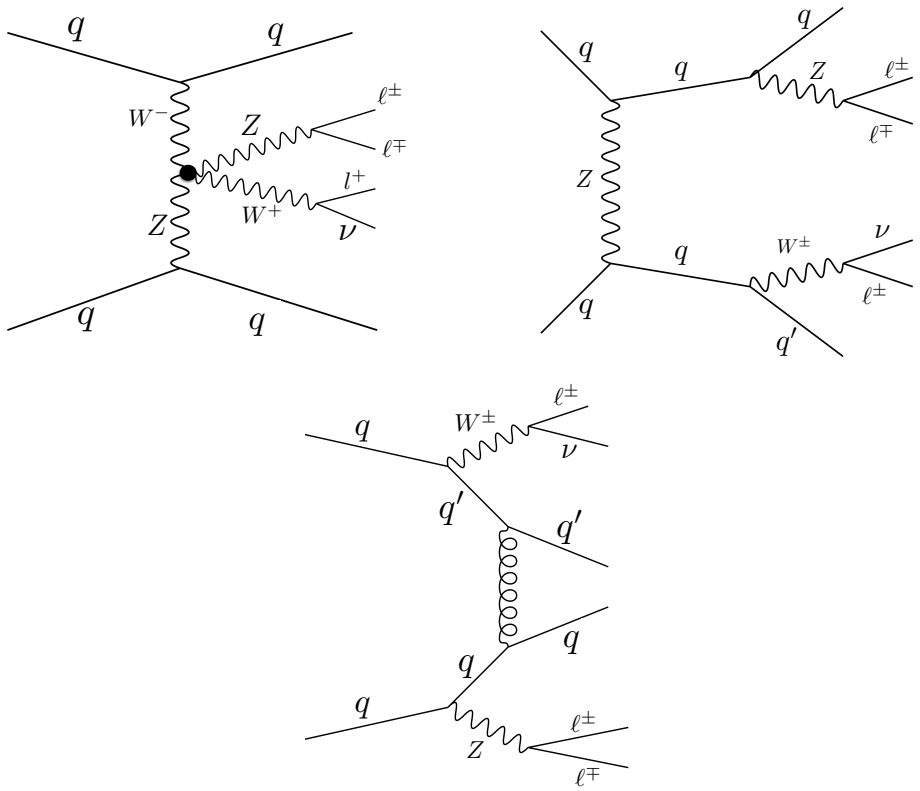

Figure 2.9: Representative tree level diagrams for the electroweak (the top two diagrams where the left one refers to WZ scattering) and QCD production (the bottom) of W and Z bosons with two jets.

quartic coupling interactions.

The VBS process refers to the electroweak production of vector bosons where they scatter (interact) with each other, while the electroweak production of the vector bosons without the scattering of the vector bosons with two jets have the same final state with similar topology. Therefore it is not fully possible to discriminate the scattering processes from the non-scattering ones. The corresponding electroweak and QCD production diagrams for the W and Z bosons with 2 jets are shown in Fig. 2.9, where the left plot refers to the scattering process and the right one to the non-scattering process. Therefore the VBS process is experimentally defined as the electroweak production of the vector bosons where the event selection is dominated by the scattering events.

The WZ scattering has not yet been discovered with the current data at the LHC. As a part of the work done in this thesis, the WZ scattering with 3 lepton events is studied on the prospect of the High Luminosity LHC. This study is described in detail in Chapter 8.

Part I

Measurement of the cross section for top quark pair production in association with a Z boson at 13 TeV with the CMS detector

Chapter 3

LHC and CMS

Studying elementary particle interactions directly can mainly be done with cosmic rays or in particle accelerators. The composition of the cosmic rays are mostly 90% protons, 9% alpha particles and 1% heavier nuclei (all the way up to uranium). Their interaction with the atmosphere creates secondary particles which can be detected and studied by the particle detectors on earth. The cosmic ray energies vary between 1 GeV to 10^8 TeV which covers a higher energy range compared to the particle colliders of the era, but the rate of these particles decreases with the energy and the initial energy of these particles are not known. For a brief summary of the cosmic ray studies please refer to [53]. While in laboratories, the energy, the rate and the type of the interacting particles can be controlled which makes it a very powerful way for particle discoveries and studying their properties. This can be done by sending particles to a fixed target or colliding two beams. These two have its own purposes and advantages while in collider experiments it is possible to reach higher energies and consequently more energetic final states. In fixed target experiments the collisions are asymmetric and they have higher luminosities but they produce less energetic particles.

In this chapter we introduce particle colliders which are tools to probe fundamental properties of matter and their interactions. One can think of colliders as very powerful microscopes that allow us to look closer into the structure of matter by colliding them and extracting information from their tracks or signatures. This can be done either using linear or circular accelerators and by colliding leptons, hadrons or ions depending on the purpose of the experiment. Accelerators can also be used for medical purposes, like demolishing cancer tumours or for developing better materials, which is beyond the scope of this thesis.

3.1 Colliders

In order to understand collider physics one should know the important parameters of an accelerator. The maximum energy obtainable is related to the size of an accelerator. Particles are accelerated and gain energy as they travel through the accelerators. Particles need to be charged, in order to be accelerated and focused by electromagnetic devices, and in general they need to be stable. Depending on the type of the particle, colliders have different functionalities. Electron-positron colliders have clean final states with known fixed energy, which makes them better for precise measurements. In case of hadron colliders the initial energy is not fixed, since the collision happens between the constituents of the protons quarks and gluons, that the energy varies depending on the momentum carried by the colliding partons. Therefore in hadron colliders various energy regimes can be scanned, that makes hadron colliders a discovery machine. Compared to the electron colliders it is possible to obtain higher energies for the same radius of the collider since hadrons are 200 times heavier than electrons and they lose much less energy due to synchrotron radiation. On the contrary they have a more crowded final state due to many initial and final state parton interactions.

We can have a look at some of the high energy particle accelerators noting that hadron colliders lead to discoveries while lepton colliders provide precision measurements. The W and Z bosons were discovered in UA1 experiment [54] at CERN which was a proton anti-proton collider operated between 1980 and 1990. Afterwards, the Large Electron Positron Collider (LEP) [55] at CERN operated from 1990s till 2000 at energies up to 209 GeV which accomplished precise measurements of the parameters of the Standard Model. Then the next generation hadron collider was Tevatron, a circular particle collider at the Fermi National Accelerator Laboratory. Protons with anti-protons were collided at a center of mass energy of around 2 TeV, which lead to the discovery of the top quark and to its mass measurement with 1% precision [56] and as well confirmed several predictions of the SM. Lastly there is the Large Hadron Collider (LHC) , which is the world's most powerful accelerator that will be explained in detail in the next section.

3.2 Large Hadron Collider

The idea of a new collider was already proposed in the 1980s while LEP collider was being built. The maximum obtainable energy, 209 GeV, in LEP was limited due to synchrotron radiation. Therefore in order to obtain higher energies and serve as a discovery machine, in 1994 the new collider project was approved as LHC at CERN. It was decided to be built in the 27 km LEP tunnel to be able to use the existing facilities from LEP and as well to reduce the cost of the project. The design of the LHC [57] enables proton-proton

and also heavy ion collisions. LHC has been in operation since 2010 at proton-proton collisions at center of mass energies of 7, 8 and 13 TeV, while it has a maximum design energy of 14 TeV. The design of the LHC [57] ables proton-proton and also heavy ion collisions. The obtainable energy in LHC is constrained by the circumference of the tunnel and by state-of-the-art of the technology. LHC is a succession of many machines and acceleration of the protons happen in several steps that will be briefly explained. Note that the LHC machine parameters evolved over the past years and here we will summarise the status for 2016 data taking period.

Protons are first obtained by stripping electrons from the hydrogen atoms, by injecting hydrogen gas into a metal cylinder and applying electrical field to separate the protons and electrons. Then they are injected into a linear accelerator Linac2[1] where their energy reach 50 MeV. Subsequently, protons go into the Proton Synchrotron Booster, where they are accelerated to 1.4 GeV and sent to Proton Synchrotron where they reach an energy of 25 GeV. Lastly they are sent to the Super Proton Synchrotron and ready to be transferred to the LHC at 450 GeV energy. The chain of the particle accelerators can be seen in Fig. 3.1. In the main LHC tunnel, protons are accelerated for 20 minutes to reach to 6.5 TeV.

The LHC uses radio frequency (RF) cavities to accelerate protons. RF cavity is a metallic chamber formed in a specific shape and size, placed at intervals along the accelerator. The applied electromagnetic field becomes resonant and increases in the cavity where protons feel the overall force and direction of the resulting electromagnetic field and gain energy in each circulation. The arrival time of the protons is crucial. The field in the cavities oscillates at a given frequency, if the arrival of the protons has a different frequency than the RF then they would be decelerated. As a consequence of this, proton beam needs to have a bunch structure and the frequency of the bunches is synchronised with the RF frequency. When protons reach their maximum energies, RF cavities keep the protons in bunches. The LHC uses eight cavities per beam, each delivering 2 MV at 40 MHz, as a consequence these bunches are injected at every 25 ns. Under nominal conditions each proton beam has 2808 bunches with each bunch consisting of 10^{11} protons.

Protons circulate in the vacuumed beam pipes, to avoid collisions with gas molecules inside the accelerator. They circulate both in clockwise and anticlockwise directions almost at the speed of light. More than 50 types of magnets are needed to optimise the beam trajectory, to bend and to keep the beam focused. All the magnets in the LHC are electromagnets, where the strength of the magnetic field is proportional to the the amount of electric current and that it is, as well, possible to change the polarity of the

[1]The linac tank is a multi-chamber resonant cavity tuned to a specific frequency which creates potential differences in the cavities that accelerate the particle up to 50 MeV.

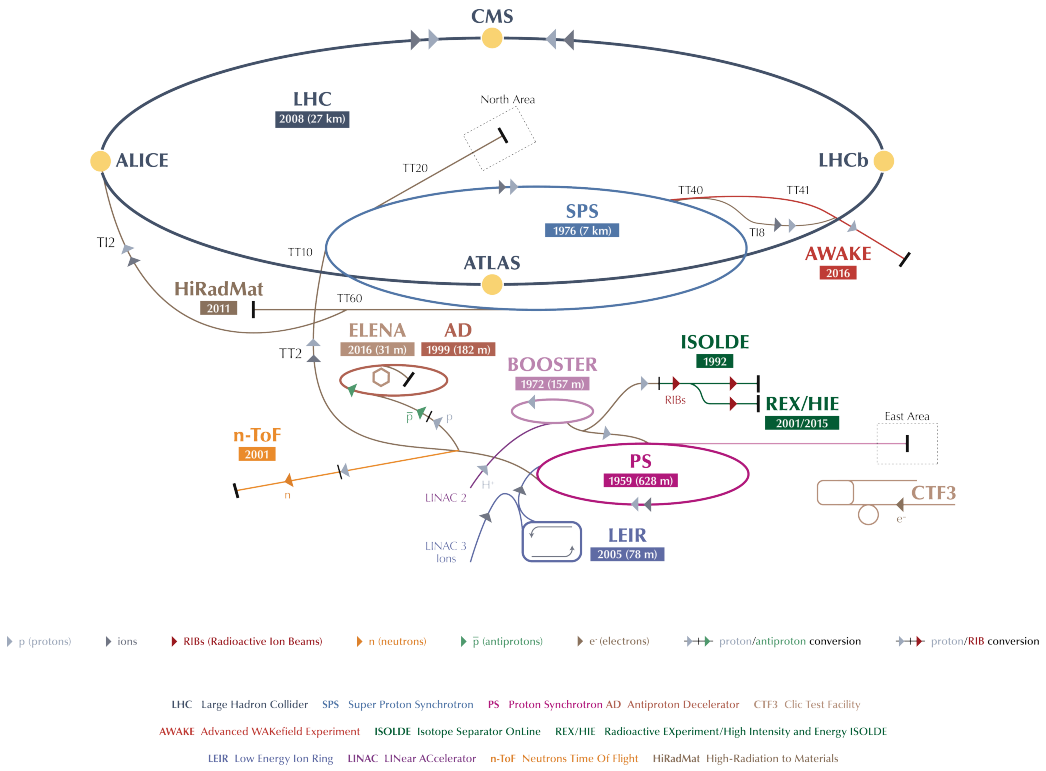

Figure 3.1: The different accelerators are shown. LHC is the last ring (dark blue line) in a complex chain of particle accelerators. The smaller machines are used in a chain to help boost the particles to their final energies and provide beams to a whole set of smaller experiments.

magnets by reversing the direction of the current. The main dipole magnets, which are needed to bend the proton beam, generate 8.3 tesla which is 100000 times more powerful than the Earth's magnetic field. In order to reach this strength, superconductor materials are used. These materials lose their resistance when they are cooled down enough which allows high currents without losing energy due to electrical resistance. The temperature where they become superconductive is below 10 K $(-263.2°)$. Still this temperature is not cold enough for the magnets to generate the required high magnetic field needed to bend 6.5 TeV beams around the 27-km ring. The optimal operation temperature of the LHC is 1.9 K $(-271.3°)$ which makes the LHC as the largest cryogenic system and also one of the coldest places on Earth. The other type of magnets used are quadruple magnets. They are used for focusing the proton beam, which adjust the beam width and height to keep it inside the vacuumed pipes. There are other multipole magnets to help the beam to focus and also to squeeze the beams at the collision points in order to increase the interaction probability.

The quantity that characterises the number of interactions is luminosity. In-

stantaneous luminosity ($\mathcal{L}$) is defined as the number of interactions per unit area per unit time, and it is one of the most relevant parameters for experimentalists. The larger is the value of luminosity, the larger is the number of collisions. In the case of two colliding beams, $\mathcal{L}$ can be expressed with the intensity of the colliding bunches (N_i), number of bunches in one beam (n_b), revolution frequency (f) and effective cross section of the colliding beams: when assuming a Gaussian beam distribution it is $4\pi\sigma_x\sigma_y$ where σ_x, σ_y are the horizontal and vertical width of the beams respectively [58]. The luminosity in terms of these quantities are defined as:

$$\mathcal{L} = \frac{f N_1 N_2 n_b}{4\pi\sigma_x\sigma_y}. \tag{3.1}$$

To calculate the number of collisions we need to consider the integrated luminosity which is the total luminosity collected over a time period and expressed in units of inverse barn b^{-1} and the interaction probability or the interaction cross section (σ). The number of collision events are:

$$N_{\text{events}} = \sigma \int \mathcal{L} dt. \tag{3.2}$$

The LHC has a design luminosity of 10^{34} cm^2 s^{-1} and the inelastic proton proton cross section at $\sqrt{s} = 13$ TeV is around 70 mb. This leads to an average event rate of 700 MHz. The total luminosity measured and recorded in 2016 by one of the main purpose detectors at CERN is shown in Fig. 3.2.

3.3 LHC Detectors

There are four big experiments installed in underground caverns of the LHC, where the particle collisions and their detection take place.

- A Large Ion Collider Experiment (ALICE) [59] is specialised to measure lead-lead ion collisions. Its main purpose is to study the properties of quark gluon plasma (QGP), a state of matter where quarks and gluons are no longer confided in hadrons at very high temperatures and densities. During the lead-lead collisions, the temperature rises up to 100,000 times hotter than the centre of the Sun. This is recreating the conditions to study QGP in the laboratory similar to those just after the big bang. The existence of such a phase and its properties are key issues in the theory of QCD and for understanding the phenomenon of confinement.
- A Toroidal LHC ApparatuS (ATLAS) [60] is a general purpose detector that was designed to primarily search for the Higgs boson and particles predicted by the BSM, as well as precise measurements of the predictions of the SM.

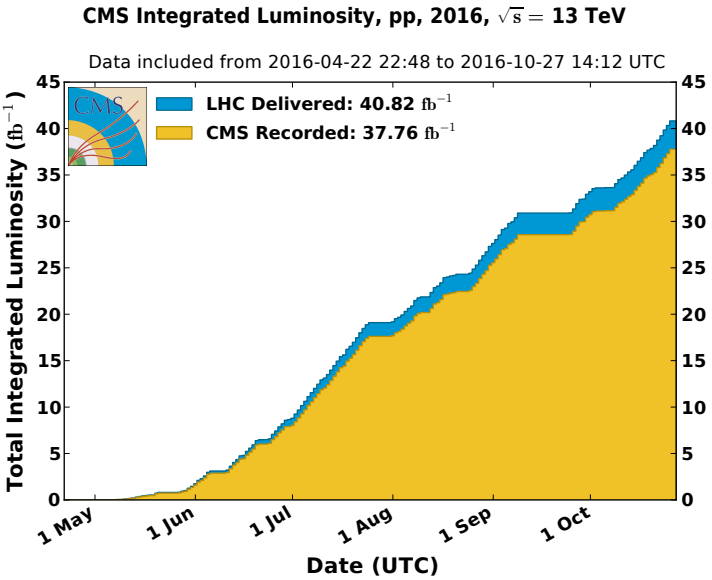

Figure 3.2: Cumulative luminosity measured online versus day delivered to (blue), and recorded by CMS (orange) during stable beams and for p-p collisions at 13 TeV centre-of-mass energy in 2016. The delivered luminosity accounts for the luminosity delivered from the start of stable beams until the LHC requests CMS to turn off the sensitive detectors to allow a beam dump or beam studies.

- The Compact Muon Solenoid (CMS) [61] is a general purpose detector with very similar physics goals as ATLAS, but it has different technical properties and a different design. The CMS detector will be explained in detail in the next sections.
- The Large Hadron Collider beauty (LHCb) [62] experiment is specifically made to study the b quark to understand the slight differences between matter and antimatter. The design of the LHCb is rather different than the general purpose detectors. Instead of surrounding the collision point symmetrically with an enclosed detector, it uses a series of subdetectors to detect mainly forward particles.

There are also three smaller experiments installed in LHC. These are TOTEM and LHCf, which focus on forward particles that emerge from collisions at small angles. MoEDAL experiment is situated near LHCb detector to search for a hypothetical particle called the magnetic monopole.

3.4 Compact Muon Solenoid

The CMS is a general purpose particle detector which is designed to detect rare and high energetic particle interactions. Like the other big detectors at the LHC, it has a layered structure of sub-detectors. The detector design

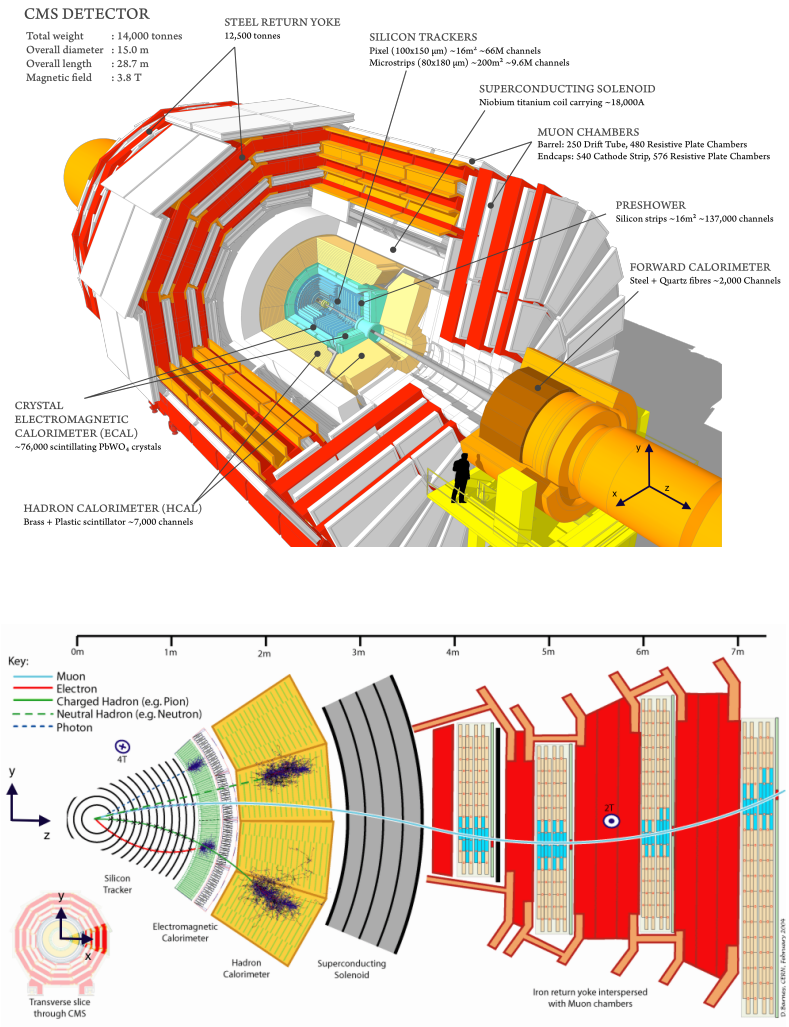

Figure 3.3: Top figure: schematic view of the CMS detector where all the sub-detectors are shown. Bottom:Transverse slice through the CMS detector, showing the individual detector subsystems and particle signatures in each. The particle type can be inferred by combining the detector response in the different subdetectors.

aims exploiting different kinds of particles and measuring their energy, momentum and charges. The main components of a general purpose particle detector are tracking devices, calorimeters, magnet system and muon detectors. The CMS sub-detectors are structured according to the geometry of the magnet system. In the central part, called as the barrel, trackers and calorimeters are embedded inside the solenoid magnet and muon systems in the iron yoke that surrounds the solenoid magnet and guides the field. The region which closes the barrel, in both sides of the barrel, is called as the forward region and end caps. The muon systems are placed in the iron yokes and forward calorimeters at the outside of the end caps.

The structure of the CMS detector showing the sub-systems and the transverse slice of the CMS with particle signatures in each sub-system are shown in Fig. 3.3. The interaction of particles in the CMS detector is briefly explained in the following. A particle emerging from the proton proton collision is firstly detected by the tracker where its trajectory and position are precisely measured. Charged particle tracks will be bended due to the magnetic field and curvature of their trajectory will give their momentum. The next layer is the electromagnetic calorimeter (ECAL), where energies of the particles that interact electromagnetically are measured and the strongly interacting particles mostly lose their energies in the second calorimeter, hadronic calorimeter (HCAL). Only muons and weakly interacting particles like neutrinos will travel beyond the HCAL. Muons will be tracked in the further muon detector systems, but neutrinos or non-interacting neutral particles will escape detection. But then by adding up the momenta of all the detected particles and assigning the total missing momentum to these particle, the transverse component of their momentum and positions can be determined. The design of the subsystems of CMS will be explained in the next sections

3.4.1 CMS Coordinate System

Before giving the details of the sub-detectors, lets introduce the coordinate system and the variables used to describe the position of particles. CMS uses a right handed cartesian coordinate system where the origin is at the nominal collision point inside CMS, the x axis is pointing to the centre of the LHC, y axis pointing up perpendicular to the LHC plane and z axis along the anti-clockwise beam direction. These coordinate positions are indicated in Fig. 3.3. The polar angle (θ) is measured from the positive z-axis and the azimuthal angle (φ) is measured from the positive x-axis in the x-y plane. Another variable is pseudorapidity (η) which gives the spatial coordinate describing the angle of a particle relative to the beam axis. It is Lorentz invariant and described as

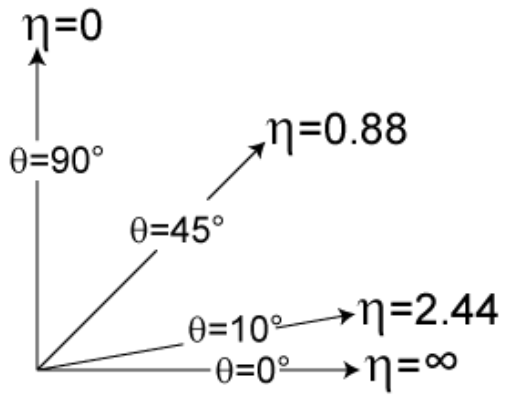

Figure 3.4: Polar angle 90° corresponds to $\eta = 0$ and pseudorapidity tends to be infinity as polar angle approaches zero

$$\eta = -\ln\left[\tan(\frac{\theta}{2})\right] . \tag{3.3}$$

The corresponding pseudorapidity for polar angles is shown in Fig. 3.4, $\theta = 90$ corresponds to $\eta = 0$ and $\theta = 10$ corresponds to $\eta = 2.5$. The central part of the detector is defined with $|\eta| < 2.5$ and the forward region for $|\eta| > 2.5$.

In order to express the angular separation between two particles a Lorentz invariant variable is used

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}, \tag{3.4}$$

where $\Delta\eta$($\Delta\varphi$) is the difference in pseudorapidity(azimuthal angle) of the two particles.

The momentum of a particle can be decomposed into a longitudinal, parallel to the beam axis (p_z), and a transverse, perpendicular to the beam axis in the x-y plane ($p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}$), components. In proton proton collisions the collision happens between the partons where the fraction of the proton's initial momentum carried by partons are not known in the longitudinal direction. While the momentum in the transverse plane (p_{T}) is known and it is initially zero. Therefore it is more convenient to use the transverse component of the momentum. Four-momentum of a particle can be expressed using p_T, η, φ together with its mass or energy.

3.4.2 CMS Detectors

The sub-systems of the CMS detector are described starting from the first detector from the collision point and then ordered by distance.

Tracker

The tracking detectors are dedicated to record charged particle trajectories and this provides information about the path of a particle. The momentum and the charge of a particle can be measured If there is magnetic field applied . At the design luminosity of the LHC there are around 1000 particles at each bunch crossing ($\sim$ 25 ns) traversing the tracker. This requires high-technology detectors with high granularity and fast response. The intense particle flux will cause severe radiation damage and the tracker system need to be able to operate in this environment for around 10 years. In addition particles should minimally interact with the detector material so that the multiple scattering, bremsstrahlung, photon conversion and nuclear interactions are avoided as much as possible. These requirements lead to the design of a tracker with semiconductor detector technology.

The CMS tracker [63, 64] surrounds the interaction point and has a length of 5.8 m and a diameter of 2.5 m. A magnetic field of 3.8 T is applied homogeneously over the full volume of the tracker. It is made of silicon pixels around the interaction point where the particle intensity is very high and of silicon micro-strips that enclose the pixels. The pixel tracker has three 53.3 cm long barrel layers and two end cap disks on each side of the barrel. This section of the tracker and corresponing layers are denoted by "PIXEL" in Fig. 3.5. The pixel sizes are driven by the desired resolution of the impact parameter, which is 100μm by 150μm. There are 66 million pixels that covers an area of 1 m^2 and the pixel occupancy is of the order of 10^{-4} per pixel and per LHC bunch crossing. The pixel tracker provides three high precision points on each charged track trajectory.

The region after the pixel between radii 20 cm and 116 cm is occupied by the silicon pixel tracker. It is composed of three sub layers. The Tracker Inner Barrel and Disks (TIB/TID) extend in radius towards 55cm and are composed of 4 barrel layers, supplemented by 3 disks at each end. TIB/TID delivers up to 4 $r-\varphi$ measurements on a particles trajectory. The TIB/TID is surrounded by the Tracker Outer Barrel (TOB). It has an outer radius of 116 cm and consists of 6 barrel layers providing another 6 $r-\varphi$ measurements. The TOB extends in z between $\pm$118cm. Beyond this z range the Tracker EndCaps (TEC) cover the region till $z = 282$ cm. Each TEC is composed of 9 disks, carrying up to 7 rings of silicon micro-strip detectors, they provide up to 9 $r-\varphi$ measurements per trajectory. The CMS tracker layout is shown in Fig. 3.5.

The tracker provides transverse momentum and impact parameter measurements and the corresponding resolution of these measurements as a function of η for single muons are shown in Fig. 3.6.

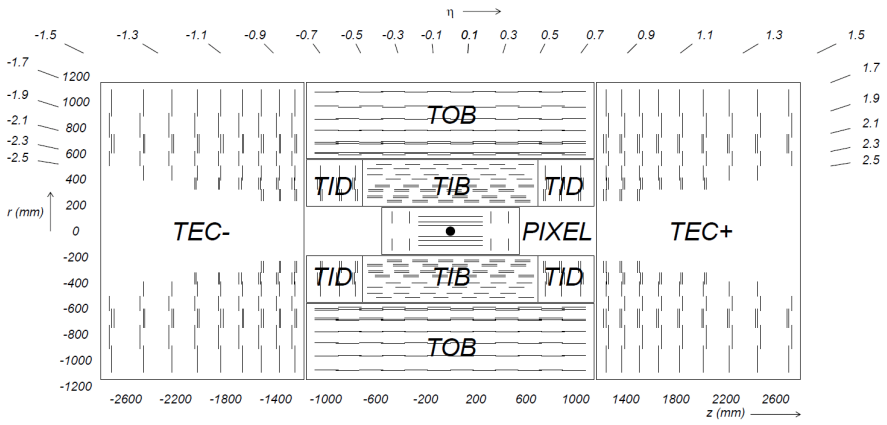

Figure 3.5: Schematic cross section through the CMS tracker. Each line represents a detector module [65].

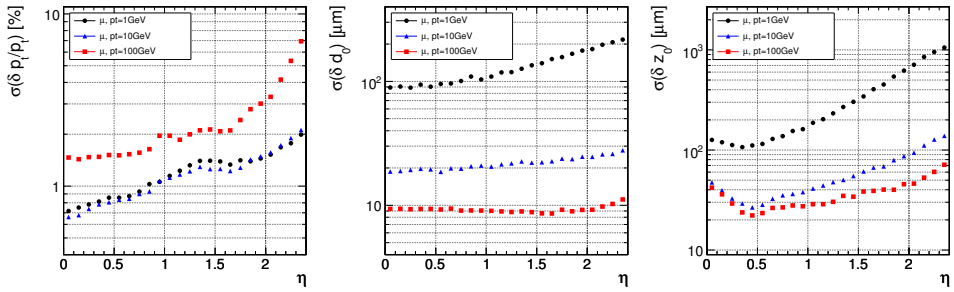

Figure 3.6: Resolution of several track parameters for single muons with transverse momenta of 1, 10 and 100 GeV: transverse momentum (left panel), transverse impact parameter (middle panel), and longitudinal impact parameter (right panel) [65].

Calorimeters

Calorimeters measure particle's energy and position. As compared to the tracking detectors which minimally interact with particles, calorimeters completely absorb the energy of a particle and therefore stop them in the detector. Calorimeters can be electromagnetic, to measure electrons, positrons and photons through their electromagnetic interactions and hadronic that are used to measure hadrons through their strong and electromagnetic interactions.

Electromagnetic Calorimeter

The CMS electromagnetic calorimeter [66, 67] played an essential role for the Higgs boson discovery. In Higgs searches in the $H \to \gamma\gamma, ZZ, W^+W^-$ decay modes, the measurement of the energy of the photons, electrons and positrons (from the decay of W and Z bosons) with good resolution is es-

sential. Since the width of the Higgs bosons with mass of 125 GeV is less than 100 MeV, the width of any observed peak would be entirely dominated by instrumental mass resolution. In addition to the good resolution, the calorimeter should be able to perform under LHC conditions: a high magnetic field, high levels of radiation and high particle intensity. In order to sustain these conditions lead tungstate ($PbWO_4$) crystals were chosen based on the fact that they produce light very fast and generate short well-defined light bursts allowing a fast and precise measurement.

The ECAL is composed of a barrel, surrounding the tracker, and two endcap sections. The barrel part of the ECAL covers the pseudorapidity range $|\eta| < 1.479$ and the end-caps cover the rapidity range $1.479 < |\eta| < 3.0$. The cylindrical barrel consists of 61200 crystals formed into 36 supermodules and the ECAL endcaps seal off the barrel at either end and are made up of 15000 further crystals. The crystal length is 23 cm in the barrel and 22 cm in the endocarp which corresponds to around 25 radiation lengths. Lead tungstate has a very high density 8.28 g cm^{-3} and short radiation length 0.89 cm. When electrons or photons pass through the ECAL they collide with the nuclei of the crystals and deposit their energy by bremsstrahlung or pair production generating a scintillation light in the crystals. The scintillation light is collected and attenuated by avalanche photodiodes and vacuum phototriodes in the barrel and endcaps respectively.

ECAL also contains Preshower (PS) detectors that are placed in front of the endcaps. The main function of the PS is to distinguish high energetic photons from low-energy photon pairs that could mimic high energy photon signals when the angle between the two emerging photons from the decay of a neutral pion is small enough to cause this problem. The CMS ECAL layout is shown in Fig 3.7 where the each sub components are shown.

Finally the energy resolution of the CMS ECAL was determined for seven electron energies between 20 and 250 GeV. The results were then fitted as a function of energy according to:

$$\frac{\sigma}{\mathrm{E}} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus \mathrm{C} \tag{3.5}$$

where S is the stochastic term which means intrinsic fluctuations from shower development, N is the noise term due to the electronic noise of the readout chain and C is the constant term that can be caused by instrumental effects. Their values are measured in the test beams and found to be S= 2.8%, N= 12% GeV and C= 0.3% [68].

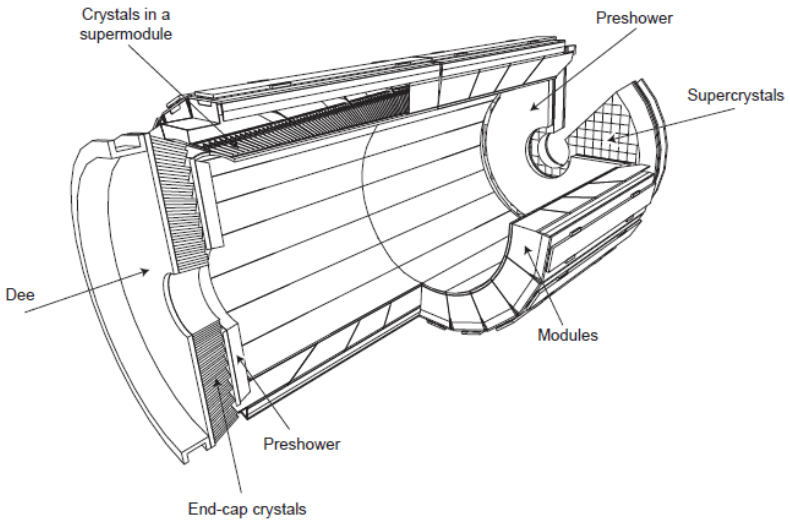

Figure 3.7: The ECAL calorimeter showing the different modules in the barrel, the endcaps and the preshower.

Hadronic Calorimeter

The CMS HCAL is designed to measure the energy of charged and neutral hadrons and indirectly energy of very weakly-interacting particles such as neutrinos. In addition it will contribute to the identification of electrons, photons and muons. It is the outer detector in the barrel completely surrounding the ECAL. It is designed to be as hermetic as possible and covers a pseudorapidity range up to $|\eta| = 5.2$. This is to ensure there are no particles escaping from the detection which is essential for the determination of the missing transverse energy. The HCAL consists of four parts central (HB) and outer barrel (HO), end caps (HE) and forward (HF). HB and HE, both embedded in the CMS magnet, are sampling calorimeters consisting of active material inserted between copper absorber plates and cover the pseudorapidity range up to $|\eta| = 1.3$ and $1.3 < |\eta| < 3.0$ respectively. The active elements of the central hadron calorimeter are 4mm thick plastic scintillator tiles. The light is collected by wavelength-shifting plastic fibers and readout by a set of hybrid photodiodes. Because of the space constraint within the magnet, the HB thickness is limited to 5.8 hadronic interaction lengths at $|\eta| = 0$ and increases to 10 interaction lengths at $|\eta| = 1.2$. In order to cover all the hadronic showers to be detected in HB and HE, layers of scintillators are placed outside the magnet which constitute the HO. About 5% of all hadrons above 100 GeV deposit energy in HO. The HF completes the hermetic structure of the HCAL covering the pseudorapidity range $3.0 < |\eta| < 5.2$ in the very forward regions and is placed z= $\pm$11.2 m from the interaction point. The HCAL energy resolution was measured with a pion test beam [69] to be:

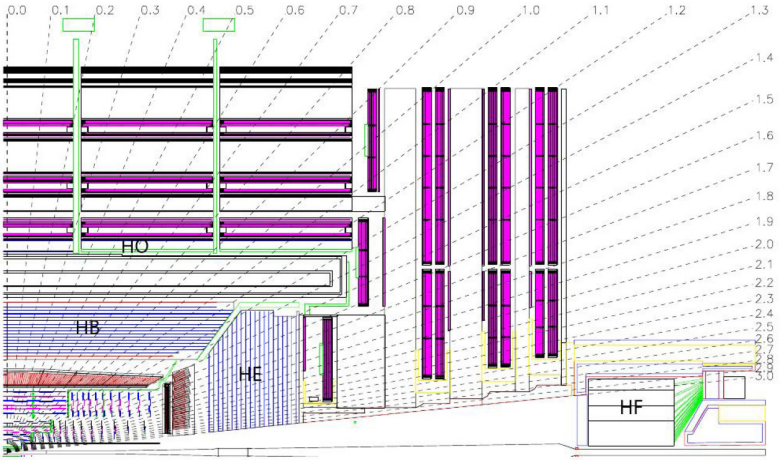

Figure 3.8: Longitudinal view of the CMS detector showing the locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters where the pseudorapidity coordinates are shown

$$\frac{\sigma}{\mathrm{E}} = \frac{85\%}{\sqrt{E}} \oplus 7\%. \tag{3.6}$$

The each sub components of the CMS HCAL layout is shown in Fig 3.8.

Muon System

As the name of the experiment implicates, a dedicated muon system with precise and robust muon measurement is one of the main tasks of CMS experiment. Muons are minimum ionising particles that they barely interact with matter and are not stopped in the calorimeters. Therefore no other detectable particle can reach the outermost muon systems rather then muons with clean signatures. This makes muons a powerful tool for recognising signatures of interesting processes over the very high background rate expected at the LHC.

The CMS muon system [70, 71] is composed of three sub systems of different gaseous particle detectors which are used for muon identification, momentum measurement and trigger. The design of the muon system, like the other inner detectors, is driven by the shape of the solenoid magnet. It has a cylindrical barrel section covering $|\eta| < 1.2$ and two planar endcap regions $1.2 < |\eta| < 2.4$. The barrel region, where the muon rate is low and the magnetic field is uniform, consists of drift tube (DT) chambers and organised in 4 stations. Each station has 2 sets of 4 chambers that provide muon coordinates in the $r - \varphi$ plane and also the z position. In the end-cap regions the muon rate are high and magnetic field is non-uniform. Therefore muon

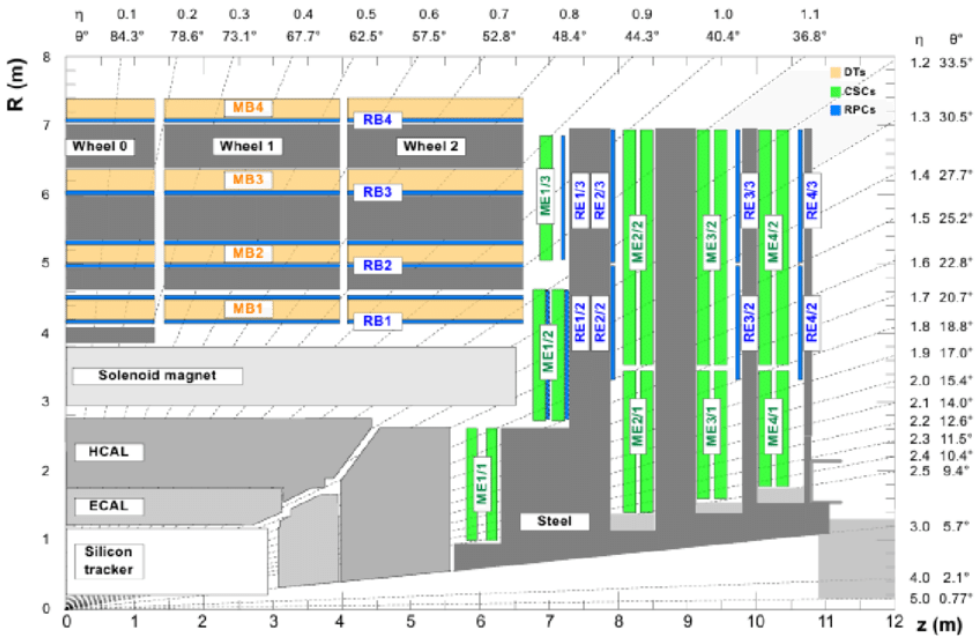

Figure 3.9: Quarter-view of CMS with labeled muon barrel (MB) and endcap (ME) stations. The steel yoke is represented by darkly shaded (red) blocks between the muon chambers. Pseudorapidities and polar angles are indicated on the top and right edges of the diagram.

system uses cathode strip chambers (CSC) with fast response time, fine segmentation and radiation resistance. CSCs identify muons in $0.9 < |\eta| < 2.4$ and are composed of 4 stations of each with 6 layers. Each chamber consists of cathode strips running radially outward and anode wires running approximately perpendicular to the strips. The cathode strips provide a position in the $r - \varphi$ plane and anode wires provide measurements of $|\eta|$ and the beam-crossing time of a muon. In addition to the precise space and time information DT and CSCs can each be used for muon trigger that will be explained in Section 3.5. Resistive Plate Chambers (RPC) are added in the muon system as a complementary dedicated trigger system. They are installed in the barrel and on each endocarps and cover the pseudorapidity range $\eta < 1.6$. The layout of the muon system is shown in Fig.3.9

3.5 Trigger and Data Acquisition

The role of the trigger and data acquisition (DAQ) system of the CMS is to decide and record collisions with the highest probability of an interesting event. As mentioned before the LHC delivers an average of 20 proton proton collisions at the nominal design luminosity of 10^{34} cm^{-2}s^{-1} every 25 ns, corresponding to 10^9 interactions every second. It is not possible to

record each interaction with current technology and resources, therefore a selection mechanism is needed in order to reduce the size and also to decide on physics wise interesting events. This task is performed by the trigger system, then this event is processed by the DAQ. The event selection decision is made in two steps called Level-1 (L1) trigger and high level trigger (HLT). The bunch crossing frequency, 40 MHz, is too short to make a decision and store megabytes of data for each event. Therefore at L1 all the data is stored at for 3.2 μs (128 bunch crossings), this period time is constrained by the the physical size of the CMS detector and underground caverns. During 3.2 μs, trigger data must be collected from the front end electronics, decisions must be made and propagated to the readout electronics front-end buffers. The events of interest from 40 MHz event rate is selected based on lookup tables since the decision is needed to be made very fast. These tables have information of all possible physics processes such as Higgs decay modes, BSM signal modes, heavy quark decays, exotic particles and also events needed for SM and soft physics. L1 reduces the event rate down to approximately 100 kHz, to a rate which is acceptable by the readout electronics.

The decision is made using the information from calorimeters and the muon system. All three detectors in the muon system participate in muon trigger system and find track candidates in each stations, layers and form trigger tracks. DT and CSCs use local and regional trigger components while RPCs directly compare the tracks with predefined patterns. Among the tracks from all detectors, the four highest quality muon candidates are forwarded to the global muon trigger along with quality bits for each track.Lastly the global muon trigger selects four muon candidates for each bunch crossing. The object constructed at trigger level are called trigger objects. The trigger decision in the calorimeter is done in a similar way. The trigger tower energy sums from ECAL and HCAL, called trigger primitives, are forwarded to the regional calorimeter trigger. ECAL, in addition to the energy sum, calculates the transverse size of the shower to distinguish electron and photons from jets. The trigger primitives are forwarded to the Regional Calorimeter Trigger where the energies from the trigger towers are summed and used to identify electrons, photons, jets and taus. Next and final step is operated in Global Calorimeter Trigger. At this stage the trigger objects are constructed and this information goes to the L1 system.

Lastly the information from muon and calorimeter trigger systems are fed into the L1 global level trigger and at this stage a L1 accept or reject decision is issued for every bunch crossing. The layout of the L1 decision is shown in Fig. 3.10.

The HLT event selection is performed in a similar way to that used in the offline processing. After an L1 accept each event is accepted and propagated to the HLT system via DAQ. For the event decision HLT has access to the

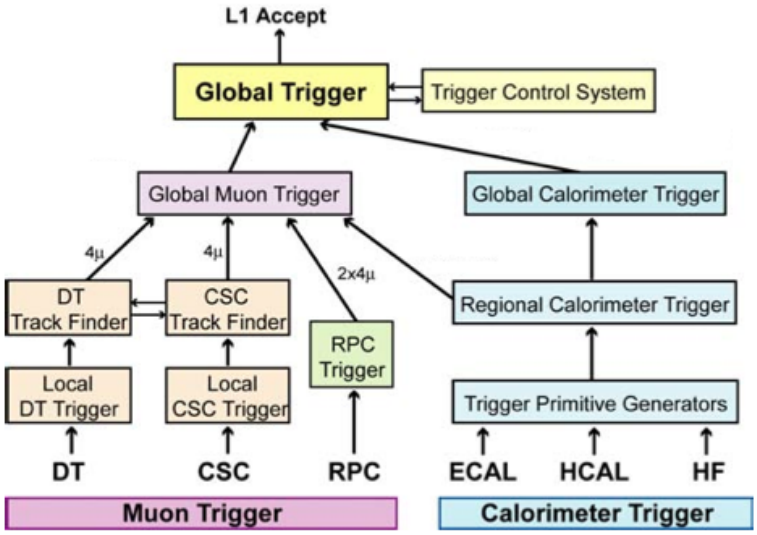

Figure 3.10: Block diagram of the CMS level 1 trigger system.

full detector readout and can perform set of algorithmic processes using reconstructed objects such as electrons, muons and jets. The HLT is composed of many trigger paths, each corresponding to a dedicated set of selection criteria such as three muons with specific transverse momentum, having a missing transverse energy above a threshold. All the specifications of a path can be changed depending on the physics needs and collision conditions. After the L1, the initial event rate is reduced to 100 kHz, and with HLT it is reduced to 1 kHZ which is the sustainable level for storage and physics analysis.

3.6 Grid Computing

The volume of data collected at the LHC, approximately 25 GB/s, requires a computing power far beyond the capacity available to CERN. It necessitated a grid computing which is the collection of computer resources from multiple locations. The national and regional computing facilities of the laboratories and universities collaborating on the LHC were integrated to a single LHC computing service, the Grid, in 2002. These services are arranged in different levels called as ”Tiers”. The data and simulation samples from the LHC experiments are distributed around the globe in levels of tiers and available to all to collaboration for the storage and analysis of the data.

Chapter 4

Event Generation

In the previous chapter the experimental tools needed to detect the experimental observables were explained. In the next chapters we will go through how to interpret and analyse what we see in the detectors. As in other areas of science, in experimental particle physics one needs to compare the experimental observables with the theoretical predictions. In this chapter we will briefly explain how to interpret the theoretical predictions. The main tool for the implementation of the theory is to use Monte Carlo (MC) event generators to simulate the event features of a collision: the signal processes and their backgrounds. Then detector simulations are needed to simulate the passage of the final state particles through the detector material.

4.1 MC Event Generation

The structure of a proton-proton collision at the LHC needs to be implemented by the MC event generators using the existing knowledge of SM and guesses on BSM. The understanding of the final state particles in proton-proton collisions is a very challenging problem. A proton-proton collision is illustrated in Fig. 4.1 and this figure will be referred for building up the simulation of a proton proton collision by MC event generators through the following steps:

1. **Hard process:** is defined by the collision of two beam constituents at a high momentum scale and consists of the most energetic final states. It is denoted as the central red blob in Fig. 4.1. The simulation of this processes, which involves large invariant momentum transfer, is the core and the first step of any simulation through Monte Carlo event generators. Thus this calculation is not very straightforward since it involves non-perturbative calculations. According to the asymptotic freedom of QCD, hadrons interact weakly at high energies corresponding to a smaller coupling constant, α_s, so that the constituents of the

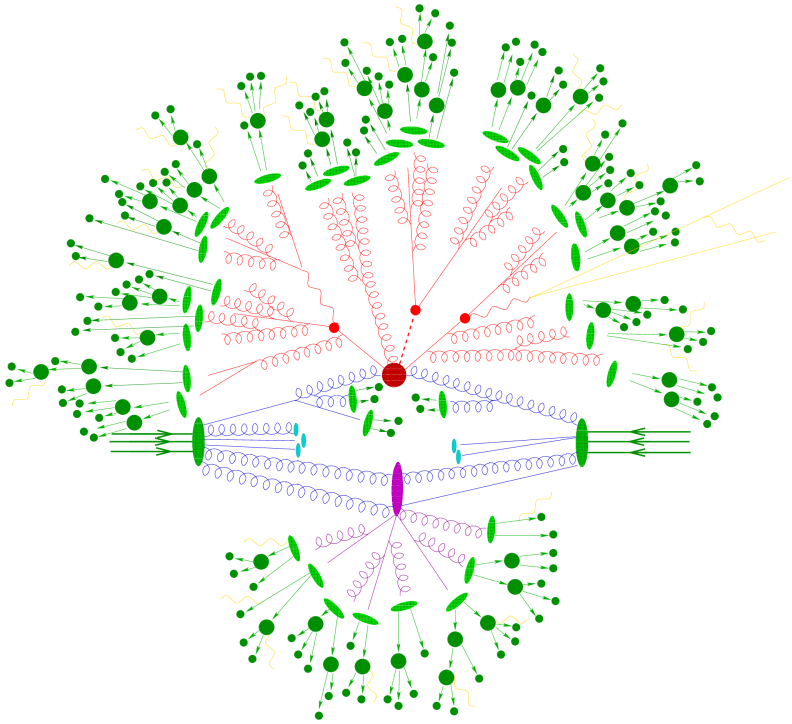

Figure 4.1: The big green blobs accompanied by the arrows illustrate protons where their constituent quarks and gluons, denoted by blue lines, interact. The red blob in the center represents the *hard collision* or *hard scattering*, surrounded by a tree-like structure representing parton showers. This process is usually the interaction of interest consisting of the most energetic final states. The purple blob indicates a secondary scattering. Parton-to-hadron transitions, hadronization, are represented by light green blobs, dark green blobs indicate hadron decays, while yellow lines signal soft photon radiation, taken from [72]

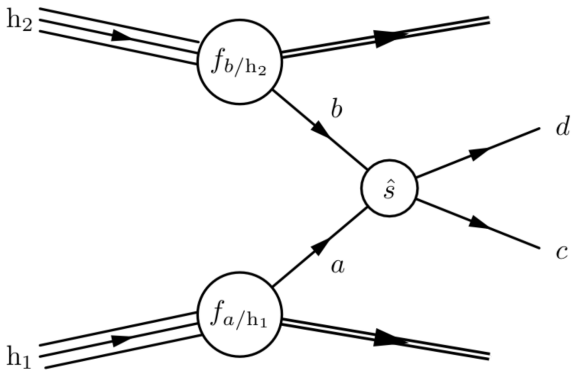

Figure 4.2: Hard scattering of parton a from hadron h_1 and parton b from hadron h_2 to partons c, d.

hadron can be regarded as free particles. Whereas, at low energies the interaction becomes stronger as the α_s becomes larger and partons confine into hadrons [73]. The high energetic interactions, also called short-distance interactions, can be calculated perturbatively while in case of low energetic, long-distance, interactions this is not possible due to large α_s. Therefore the so-called factorisation theorem brings a solution to this problem by resolving the short distance parton cross section from the long distance interactions [74]. In Fig. 4.2 the application of the factorisation theorem for partons a and b, from hadrons 1 and 2, scattering to c and d partons is illustrated. This can be expressed by the following equation:

$$d\sigma^{h_1 h_2 \to cd} = \int_0^1 dx_1 \int_0^1 dx_2 \sum_{a,b} f_{a/h_1}(x_1, \mu_F^2) f_{b/h_2}(x_2, \mu_F^2) d\hat{\sigma}^{ab \to cd}(\mu_R^2, \mu_F^2) \tag{4.1}$$

here $f_{a/h_i}(x_i, \mu_F^2)$ is the parton distribution function (PDF). The PDF $f_{a/h_i}(x_i, \mu_F^2)$ gives the probability of finding a parton of a with momentum fraction x_i of the hadron h_i at the energy scale μ_F. Here μ_F is the $factorization\ scale$, which characterises the hard scattering and can be thought as the scale that separates the long- and short-distance interactions and μ_R is the *renormalization scale*, which is a scale used to fix the divergences of loop diagrams. The PDF's can not be obtained via perturbative QCD calculations that they are computed by fitting the data from several experiments and many different processes. This is possible due to the fact that the PDF's are process independent meaning that they are universal. They can be measured in one process then can be applied to other processes. Their evolution to any scale can be calculated by DGLAP evolution functions once they are measured in one scale, detailed explanation can be found in [75]. The hard interaction differential cross section for a and b scattering to c and d is denoted by $d\hat{\sigma}^{ab \to cd}(\mu_R^2, \mu_F^2)$. This term contains only hard emissions above the factorisation scale μ_F and can be calculated by perturbative QCD.

The leading order (LO) calculations contain tree level Feynman diagrams and next-to-leading order (NLO) calculations contain diagrams with one-loop corrections[1]. Higher order next-to-next leading (NNLO) calculations involve higher order loops and corrections. The simulation samples used in this thesis are based on calculations at either LO or NLO precision.

[1]It takes into account all real emissions and virtual corrections at higher orders in α_s [76]

2. **Parton shower**: The simulation of the proton proton collision is followed by the parton shower. The final state partons carrying a color charge can emit gluons (QCD radiation) and also can interact with each other emitting further gluons. This process is called as the parton shower, denoted by the red spiral tree structure surrounding the hard interaction in the Fig. 4.1. It evolves until the partons lose energy due to gluon emission and they go into the hadronization phase.

3. **Hadronization**: In the process of partons losing their energy by QCD radiation, at some energy level the interaction among the colored partons become stronger, i.e. α_s becomes large, and they are bounded into colourless hadrons. This transition is called *hadronization*. Hadrons are the first experimental observables of the event generation in an event. The hadronization process roughly happens at an energy of 1 GeV where this energy depends on the hadronization model. The most common hadronization models are the cluster model [77] and the Lund string model [78]. For a detailed explanation of these models please see [79]. The transition of partons to hadrons are denoted as the light green blobs in Fig. 4.1.

4. **Hadron decays**: Most of the hadrons produced in the previous step are unstable and they decay further, until a set of particles is obtained that can be considered stable on time scales relevant to the given measurement[2]. These stable hadrons are the final observables detected. Therefore the decay modelling has an important impact on the final state yields and spectra. The hadron decays are shown as the dark green blobs in Fig. 4.1.

5. **Secondary interactions**: Up to now, the interactions of the partons that are not coming from the hard collisions have not been considered. At first approximation it can be assumed that these partons do not interact and just fly away undisturbed. But in reality this is not the case and partons not coming from the hard collision can also interact with each other. These interactions are called as multiparton interactions. In a proton proton collision the primary spectator partons (beam remnants) can split or emit gluons and hadronize. In addition, the initial and final state gluon radiations not connected to the hard collisions and the multiparton interactions are called as the underlying event. The illustration of a secondary interaction is denoted in purple blob in Fig. 4.1.

After all these processes the four-momentum vector for each generated particle with its unique particle identity is saved to be used when analysing

[2]E.g., a typical hadron-collider definition of a stable particle is defined by the distance the particle travels, $c\tau \geq 10$ mm, where τ is the lifetime of the particle. This includes the weakly-decaying strange hadrons (K, Λ, $\Sigma^{\pm}$, Ξ, Ω) [32, 80]

the simulated event samples. We will refer to this set of information as "generator or particle level info"

4.2 MC Event Generators

There are different kinds of event generators used in high energy physics. These are general purpose event generators such as PYTHIA [81], [82] that simulate all the stages of a proton collision mentioned in the previous section, and matrix element (ME) calculators delivering an event at the parton level such as Powheg, MadGraph MG5_aMC@NLO. The different event generators used in this thesis will be briefly explained below.

MG5_aMC@NLO

MG5_aMC@NLO [83] is a framework that is used for the computations of cross sections, the generation of hard process events by the calculation of ME for SM and BSM. Processes can be generated at LO and NLO, and the NLO accuracy in the case of QCD corrections to SM processes. For completing the event simulation after the ME calculation an interface to a program such as PYTHIA is used for parton showering, hadronization and secondary interactions. While merging the matrix element with the parton shower, the additional partons from the event generation with high multiplicity may also be coming from the additional partons created by parton showering. This results in an overlap in the phase space and called double-counting and illustrated in Fig. 4.3. The double counting is corrected with the MLM merging technique for the LO processes [84]. In this scheme, as first step the number of partons before the parton showering are determined and a cut-off value for the merging scale is chosen. Once the parton showering happens, the partons are clustered to jets using the k_t algorithm which will be explained in detail in Sec. 5.2.4. The ΔR between the parton and the jet are calculated for each hard parton (parton generated by MG5_aMC@NLO) with all the clustered jets in order of decreasing hardness(energy). The parton and the jet is considered to be matched if the ΔR between the parton and the jet is smaller than the ΔR used for the jet clustering algorithm. If there is a hard parton remaining not matched to a jet, the event is vetoed. In case of all the hard partons are matched the event is kept. In case of NLO merging the FxFx [85] method is used. It uses of the CKKW Sudakov suppression [86] and MLM event rejection, detailed explanation can be found in [85].

POWHEG

POWHEG Box [88] is a dedicated NLO event generator and uses the POWHEG method [89]. In the POWHEG method the hardest radiation is generated first, with a technique that yields only positive-weighted events using the

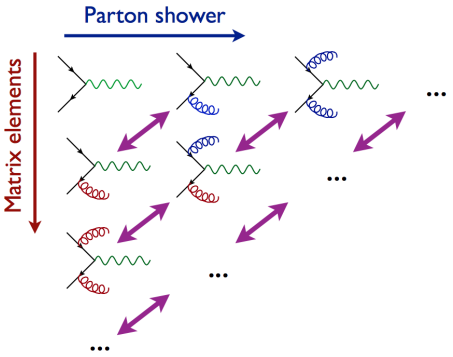

Figure 4.3: In the vertical direction the LO Feynman diagrams with increasing multiplicity are displayed in red. In the horizontal direction the corresponding diagrams with additional partons coming from the parton showering are shown in blue. The overlapping diagrams are shown by the purple arrows, taken from [87]

exact NLO matrix elements [90]. For the parton showering it needs to be interfaced with a chosen program, like Pythia.

Pythia

Pythia is a general purpose event generator. It contains theory and tools needed for the calculation of hard and soft interactions, PDFs, initial-final state parton showers, secondary interactions and also hadron and particle decays. It is commonly used for parton showering in LHC.

4.3 Detector Simulation

The last step of the simulation of a proton proton collision is the interaction of the final state particles coming from the hard interactions and hadrons coming from the parton showering with the detector. This is done by a dedicated toolkit called GEANT4 [91]. It simulates the passage of particles through matter using Monte Carlo methods. It is used by high energy physics and also used outside high energy physics such as in space applications and medical applications. GEANT4 includes all the facilities needed for the interaction of particles with the detector matter. These are listed in the following:

- detector geometry and materials;
- particle interactions in detector matter;
- tracking;
- digitisation and hit management;
- event and track management;

- visualisation;
- visualisation framework and user interface;

The listed constituents are all implemented for the CMS sub-detectors and are calibrated with the data taken from test beam, cosmic ray and collision runs. According to these calibrations, the parameters of GEANT4 are tuned. After the calibration the simulated electronic response of the particles have the same form as the data and can be used as data for event reconstruction and then can be used for physics analyses.

Chapter 5

Event Reconstruction

In this section how the information from the detectors are brought to the analyses will be described. At LHC when two proton bunches collide what we see in the detectors are digitised electronics signals of many charged particle hits, energy deposits in the calorimeters and segments in the muon systems. An example visualisation is shown in Fig 5.1. These set of information from detectors and several particle identification criteria are used to build physics objects like electrons, muons and jets. CMS adopted a so-called particle flow for this procedure.

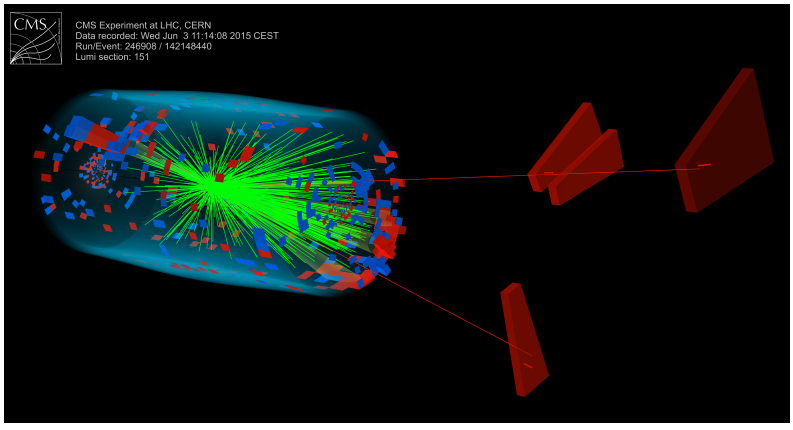

Figure 5.1: An event display of a CMS 13 TeV proton proton collision. The green lines are the charged particle tracks in the tracker, the red clusters are energy deposits in ECAL, the blue clusters are energy deposits in HCAL and red segments on the forward region corresponds to muon segments in the muon system

5.1 Particle Flow Algorithm

The Particle Flow algorithm (PF) [92] aims for reconstruction and identification of all stable particles in an event. The PF elements are particle tracks in the tracker and the muon system and energy clusters in the ECAL and the HCAL. The PF reconstruction algorithm combines the information from PF elements via a linking algorithm. Considering the high particle density environment in CMS this becomes a challenging task.

5.1.1 Particle Flow Elements

Charged particle tracks

Track reconstruction aims at measuring the direction and magnitude of charged particles momenta. The tracking algorithm method used in CMS is an iterative application of the combinatorial Kalman Filter (KF) [93]. In the first step of this iterative fitting, tracks are seeded and reconstructed with tight selections. This leads to a moderate tracking efficiency with a negligibly small fake rate. The next step is to remove hits assigned to the tracks found in the previous iteration and continuing this by loosening the track seeding criteria. With this iterative technique, charged particles with as little as three hits, a transverse momentum as small as 150 MeV and an origin up to 50 cm away from the beam axis, are reconstructed with a fake rate of the order of a per cent with an efficiency of 99.5% for isolated muons in the tracker acceptance, and larger than 90% for charged hadrons in jets [94].

Calorimeter clusters

The other element of the PF is the clustering algorithm which targets to group energy deposits in the calorimeters. The clustering is performed for each calorimeter subdetectors, ECAL and HCAL, separately. First, cluster seeds are identified with energy deposits above a threshold and with energies higher then their neighbour cells. Then topological clusters are formed by adding calorimeter cells to the seeds formed in the previous step, if they have at least one common cell side with each other and if the cell energies are above a given threshold which is two standard deviations of the electronic noise (*i.e.* $\sim$ 80 MeV in the ECAL barrel and $\sim$ 300 MeV in the ECAL end-caps and up to $\sim$800 MeV in the HCAL). Finally PF clusters are formed from the topological clusters according to the cell-cluster distance, with an iterative determination of the cluster energies and positions. These calorimeter clusters are then used in the linking algorithm for particle reconstruction.

Link Algorithm

A particle may give rise to more than one PF element in the detector. For instance an electron leaves a track in the tracker and also energy deposits in the ECAL. Once tracks and calorimeter clusters are reconstructed, they need to be linked to reconstruct particle candidates. The linking algorithm combines the PF elements tentatively and defines the distance between the two linked elements to assign a quality measure to the link. The PF blocks typically contain several elements that ensures to have simple inputs for particle reconstruction and identification algorithm.

The linking algorithm is based on extrapolating tracks (clusters) with clusters (tracks). The linking of a charged track with a calorimeter cluster is done as follows. The track is extrapolated from the last hit in the tracker to the two layers of PS; to the ECAL to a depth of expected maximum of a longitudinal electromagnetic shower; to the HCAL at a depth corresponding to a hadron shower length. The track is linked with any cluster if the calorimeter cluster overlaps with the extrapolated track and the distance between these two in (η, φ) is defined as the link distance. The linking of two calorimeter clusters is done when position of the most granular calorimeter (ECAL and PS) is within the cluster of the less granular calorimeter (HCAL and ECAL), and the link distance is defined in the same way. A charged particle in the tracker and a muon tracker in the muon system, since no other particle can give rise to these two elements, are assigned to a *global muon* if a global fit between two tracks has an acceptable χ^2 . For this case the link distance is the χ^2 value.

5.1.2 Particle Flow Reconstruction

For each PF block, the reconstruction starts with the cleanest final state which are muon candidates. Muon reconstruction is done as the first step of the PF algorithm. Each *global muon*, is tagged as a PF muon if its combined momentum is compatible within three sigma with the one obtained by the tracker. The PF elements corresponding to muons are removed from the block and then electron reconstruction and identification follows.

There are two methods for electron reconstruction; the ECAL-based and the tracker-based approach. The ECAL-based approach makes use of the ECAL clusters with energy greater then 4 GeV. The cluster energy and position are used to deduce the position of the expected inner hits in the tracker. Due to the interaction of electrons with the tracker, most of the electrons emit bremsstrahlung photons before reaching the ECAL. Therefore the performance of the method depends on the ability to reconstruct the energy of the electron and possibly emitted photons. The energy of the electron and of the possibly emitted bremsstrahlung photons in the ECAL are collected and grouped into a supercluster (SC).

In order to reconstruct electrons missed by the ECAL-based approach the tracker based approach is developed in the context of PF algorithm. When the muon track is removed from the block of PF track candidates, each track goes to a pre-selection stage. The corresponding tracks of non-radiating electrons or electrons emitting low energetic photons (called also as soft photons) can be reconstructed across the whole tracker with a well-behaved χ^2. Thus electrons radiating high energetic photons lose energy by Bremsstrahlung and have short and deviated tracks in the tracker. These tracks are fit with a Gaussian-Sum Filter (GSF) [95] to follow their trajectory to the ECAL where it can be matched with the closest ECAL cluster. The ratio of the ECAL cluster energy to the track momentum need to be compatible with unity to form an electron seed. In case of soft photon emission, the method above holds but the track fit has a larger χ^2. If more energetic photons are radiated, the fit may not be able to accommodate the difference in electron's momentum and causes the track to be reconstructed with less hits. A preselection based on the number of hits and the fit χ^2 is applied and the selected tracks are refitted with GSF. The electron seeds obtained by the ECAL and tracker-based methods are submitted to the full electron tracking.

In order to select tracks for the reconstruction of hadrons, tighter quality criteria are applied to the remaining tracks. If the relative uncertainty of the measured track is larger than the expected relative calorimetric energy resolution for charged hadrons, these tracks are removed from the remaining PF collection.

Hadrons are the remaining particles to be reconstructed in PF algorithm. The ECAL and HCAL clusters not associated to any tracks within the tracker acceptance $|\eta| < 2.5$ can be associated to photons and neutral hadrons. The energy deposit in the ECAL is associated to photons and the energy in the HCAL to neutral hadrons because in hadronic jets 25% of the jet energy is carried by photons while neutral hadrons leave 3% of their energy in the ECAL.

Beyond the tracker acceptance, since there is no information on the charged tracks, charged and neutral hadrons can not be distinguished and the precedence given to photons does not hold anymore. Therefore in this region, the ECAL clusters linked to the HCAL clusters are assigned to hadrons and ECAL clusters without such a link are assigned to photons. The energy of these photons and hadrons are then calibrated using the energy values in different subsections of the calorimeters. The details of this calibration procedure is explained in [92]. After the calibration, if there is a calorimeter energy excess with respect to the sum of the associated charged momenta by an amount larger than the expected calorimetric energy resolution for hadrons, this excess is associated with photons or neutral hadrons. If the

excess is smaller than the ECAL energy and larger than 500 MeV it is identified as a photon. The energy corresponding to this excess is recalibrated under the photon hypothesis and assigned to the photon energy. If the excess is large than 1 GeV then it is identified to a neutral hadron. In case of no excess, it is assigned to a charged hadron. In rare cases the calibrated calorimetric energy is significantly smaller than the sum of the track momenta. If the difference is larger than three standard deviations, then the PF muon algorithm is revisited.

5.2 Physics Objects

5.2.1 Muon Identification

The inner tracker provides precise measurements of the muon momenta. This together with the information from the muon systems leads to a very high muon reconstruction and identification efficiency. The muon object collection is composed of three different muon types depending on which sub-detectors are included in the reconstruction:

- **Muon detectors:** Hits in the DT and CSC are clustered to form track segments and used as seeds for the iterative fitting to reconstruct the muon track using the all DT, CSC and RPC hits. The final fitted track is called as *standalone muon track*.
- **Tracker and muon detectors:** As explained in the linking algorithm, a $standalone\ muon - track$ is matched to a track in the inner tracker and is fit to from a *global muon track*.
- **Tracker only:** All tracks with transverse momentum greater than 0.5 GeV and a total momentum in excess of 2.5 GeV that has at least one extrapolated match in the muon system is called a *tracker* muon.

Prior to any selection, charged hadrons may be mis-reconstructed as muons e.g. if some of the hadron shower remnants reach the muon system (punch-through). Therefore additional selections based on the quality of the muon track (this will include both global and tracker) are applied to identify muons with high efficiency and low misidentification. The main sources of muons vary depending on the momentum range of the muon. For those with transverse momenta between 10 and 30 GeV, the main sources of muons are b, c quarks and muonic K and π decays. These muons are usually produced in jets and accompanied with other particles. The muons with higher momenta usually come from the W and the Z boson decays, called prompt muons, and are accompanied by particles from hadronic activity, underlying event and pile-up [96]. Therefore isolation is a way to determine whether the lepton (applied also in electrons) originates from a boson or hadrons. A muon can be considered as isolated if the sum of the transverse momenta

of the tracks and of the transverse energy deposits with a distance of ΔR to the muon direction is required to be less then 10% of the muon transverse momentum. The isolation criteria is sufficient to reject hadrons that would be misidentified as muons therefore to a global isolated muon no further selection is applied.

After global isolated muons have been selected, depending on the analysis needs other selections are applied to the remaining ones. The muon selections used relevant for this thesis are discussed in detail in Sec. 6.3.1.

5.2.2 Electron Identification

Electrons often interact in the tracker and emit bremsstrahlung photons, and photons convert into e^+ e^- pairs which can again lead to bremsstrahlung photons. Therefore electron and isolated photon identifications are performed in a similar way. In the PF algorithm, an electron candidate is seeded from a so-called GSF track provided that the corresponding ECAL cluster is linked to a maximum two additional tracks. The photon candidates are seeded from the ECAL supercluster with an energy greater than 10 GeV and with no corresponding GSF track. For ECAL-based electrons and photon candidates the energy measured in the HCAL should not exceed 10% of the ECAL supercluster energy. Electron identification needs further criteria. In CMS, there are two kinds of electron identification. These are called as cut based electron Id or Multi Variate Analysis (MVA) based electron Id. For the cut based electron id, electron identification variables are applied individually and for the latter one they are combined in an MVA tool [97] and trained separately for prompt and non-prompt electrons. In this thesis, the MVA electron id is used since it was found to be more efficient in reducing the non-prompt lepton background and also it gave a better signal sensitivity over background. The input variables used for the MVA identification classifier are composed of:

- cluster shape variables
- tracking variables
- track-cluster matching variables

The detailed overview of these variables are given in Tab. 5.1.

5.2.3 Primary Vertices

As mentioned before, when two proton beams collide multiple proton proton collisions occur each of which leads to an interaction point, called primary vertex. The preselected tracks obtained by KF are clustered along the beam line and a preselection, based on the impact parameter and transverse momentum, is applied for primary vertex identification. Vertex candidates

Observable type	Observable	Definition
cluster shape	$\sigma_{i\eta i\eta}$, $\sigma_{i\varphi i\varphi}$	Standard deviation of the energy distribution in the cluster along the η, φ direction of the energy-crystal number spectrum along η and φ
	$\Delta\eta_{SC}$	Super cluster width along η
	$\Delta\varphi_{SC}$	Super cluster width along φ
	H/E	Ratio of the hadronic energy behind the electron supercluster to the supercluster energy
	$(E_{5\times5} - E_{5\times1})/E_{5\times5}$	Circularity: the energy sums $E_{i\times j}$ of the i crystals in φ and j crystals in η centred on the seed crystal
	$R_9 = E_{3\times3}/E_{SC}$	Ratio of the energy in a 3×3 (9 crystal) cluster around the seed over the SC energy
	E_{PS}/E_{raw}	For endcap training bins only: energy fraction in pre-shower over the raw SC energy
tracking	$f_{brem} = 1 - p_{out}/p_{in}$	Fractional momentum loss as measured by the GSF fit. The momenta p_{in} and p_{out} are extrapolations of the GSF track to the vertex and ECAL respectively.
	N_{KF}	Number of hits of the Kalman Filter track of the iterative combinatorial track finder, if any
	χ^2_{KF}	Reduced χ^2 of the KF track
track-cluster matching	E_{SC}/p_{in}	Ratio of the SC energy and the track momentum at the innermost hit
	E_{ele}/p_{out}	Ratio of the energy of the cluster closest to the electron track and the track momentum at the outermost hit
	$1/E_{tot} - 1/p_{in}$	Energy-momentum agreement
	$\Delta\eta_{in} = \lvert\eta_{SC} - \eta_{in}\rvert$	Distance between the energy-weighted center of the SC and the expected shower position as extrapolated from the GSF trajectory state at the vertex
	$\Delta\varphi_{in}$	Same as $\Delta\eta_{in}$ but along φ
	$\Delta\eta_{seed} = \lvert\eta_{seed} - \eta_{out}\rvert$	Distance between the pseudorapidity of the seed cluster and the expected shower position as extrapolated from the GSF trajectory state of the outermost hit

Table 5.1: Overview of input variables used for electron MVA id, taken from [98].

are fitted and filtered according to their compatibility with the beam line and goodness of the fit. Then they are ordered by the quadratic sum of the p_T of their tracks. The primary vertex with the highest $\sum p_T^2$ is defined as the main interaction or hard scatter vertex and the others as the pile-up vertices, which in the rest of the text will be referred as pile-up(PU). The resolution of the primary vertex is around 20 μm in the transverse (x, y) plane and around 30 μm along the beam axis (z) [99]. The so-called secondary vertices which are the vertices that correspond to the long lived particle decays, such as heavy mesons that contain b or c quarks, are also identified with dedicated techniques and will be explained in detail in Sec. 6.6.4.

5.2.4 Jets

Jet clustering

When discussing about hard scattering particle production in hadron colliders, one refers to quarks and gluons. Yet quarks and gluons are notdirectly observable due to the colour confinement, induced by the strong interaction which leads them to hadronise generating to a collimated spray of hadrons. This collection of hadrons is called a jet and is illustrated in Fig. 5.2.

Experimentally jet definition is not very straightforward and it consists of two parts: the jet algorithm that defines how the hadrons are grouped and the recombination scheme that defines how the momenta of the hadrons are combined. There are two main jet reconstruction algorithms which are cone and sequential [101]. In basic cone algorithms, objects are clustered in a defined cone size. The "anti-k_t" is a sequential algorithm used as default in

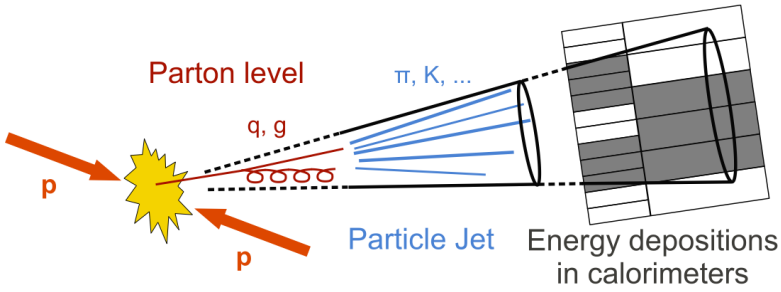

Figure 5.2: Sketch of a pp-collision and resulting collimated spray of particles, a jet [100].

the CMS experiment. It is infrared and collinear safe[1] which is not the case for the cone algorithms that were used before LHC [101]. The "anti$-k_t$" algorithm, as other sequential algorithms, is based on defining some quantity of how likely two partons are to have arisen from the same QCD splitting, and proceed sequentially to construct the jet by reconstructing the partons which are closer to each other in terms of this quantity. This is done as following. The distance between particles i and j, d_{ij}, and the distance between the entity i and the beam (B), d_{iB}, are calculated sequentially. If d_{ij} is the smallest then i and j entities are combined into a new entity k summing their four momenta, otherwise d_{iB} is assigned to a jet with index i and removed from the list of entries. The procedure is repeated until no entries left. The definition of the distances are:

$$d_{ij} = \min\left(\frac{1}{p_{Ti}^2}, \frac{1}{p_{Tj}^2}\right)\frac{\Delta R_{ij}^2}{R^2} \tag{5.1}$$

$$d_{iB} = \frac{1}{p_{Ti}^2}, \tag{5.2}$$

where ΔR_{ij} is the cone width between two particles and R is the radius of parameter where the hadrons are merged to a jet in this cone size. The functionality of the algorithm can be understood by considering several particle combinations. The distance is inversely proportional with the transverse momentum such that particles with high momentum (hard hadrons) will be given priority in the clustering and hard hadrons will tend to cluster with soft (low momenta) hadrons.

The "anti$-k_t$" algorithm uses the PF elements from all reconstructed particles except charged hadrons associated to pile-up vertices. For identification

[1] Infrared and collinear safety is the property that if one modifies an event by a collinear splitting (splitting of the hardest particle into a nearly collinear pair) or the addition of a soft emission of gluons, the set of hard jets that are found in the event should remain unchanged.

of the jets used in analyses additional quality cuts are applied, these will be explained in next Chapter.

Jet energy calibrations

The measured energy of the jets depends on several variables such as definition of the jet clustering algorithm, non-linear calorimeter response, electronic noise, pile-up, etc. All these effects lead to a difference between the measured jet energy and the energy of the jet obtained at particle level. Therefore jets need to be calibrated and corrected to attain the true particle or parton energy. In CMS, a factorised jet energy correction (JEC) procedure is applied [102]. Each level of correction adjusts different effects and scales the reconstructed jet four momentum, p_μ^{reco}, as follows:

$$p_\mu^{cor} = \mathcal{C} \cdot p_\mu^{reco} \tag{5.3}$$

where p_μ^{cor} is the corrected jet four momentum (components are indexed by μ). The overall factor $\mathcal{C}$ is composed of the offset correction C_{offset}, the MC calibration factor C_{MC}, and the residual calibrations C_{rel} and C_{abs} for the relative and absolute energy scales.

$$\mathcal{C} = C_{\text{offset}}(\rho, A, p_T^{reco}, \eta) \cdot C_{\text{MC}}(p'_T, \eta) \cdot C_{\text{rel}}(\eta) \tag{5.4}$$

Each correction is applied consequently. At first $C_{\text{offset}}(\rho, A, p_T^{reco}, \eta)$ is applied to reconstructed jet p_{T} and the p_{T} after this correction is denoted as p'_T, secondly $C_{\text{MC}}(p'_T, \eta)$ and lastly $C_{\text{rel}}(\eta)$ are applied.

- **Pile-up off-set corrections(L1):** The pile-up collisions taking place within a single beam crossing are called in-time pile up (IT PU). The contribution from the previous and subsequent beam crossings, which can also contribute to calorimetric energy in the same time window of the primary hard interaction due to the finite signal decay time in the calorimeters, are called out-of-time pile-up (OOT PU). The difference in p_T for a reconstructed jet with and without pile-up is called a pile-up offset. The corresponding correction aims to remove the energy coming from pile-up offset and also from electronic noise. It is parametrised as a function of the pile-up offset energy density ρ [103], jet area A and jet pseudo rapidity η and transverse momentum p_T. The pile-up is measured in simulation, comparing the reconstructed events with and without pile-up. The IT PU is removed by identifying and subtracting charged hadrons from pile-up vertices (CHS) and the remaining energy due to OOT PU and neutral particles is estimated per event then subtracted per jet using effective area calculation with the extended hybrid jet area method [102]. This correction is applied both to MC and data.

- **Simulated response corrections-L2L3:** The response of the CMS detectors is not uniform in terms of reconstructed jet transverse momenta and to jet pseudo rapidity. The correction factors are derived and applied on jets that are corrected for pile-up offset in MC and in data. Simulated correction factors are derived from a QCD multijet sample where a particle level jet is matched to the closest reconstructed jet if it is within half of the jet radius parameter R. The simulated particle response is defined as the ratio of the arithmetic means of matched reconstructed and particle level jets transverse momenta as a function of the particle level transverse momentum and reconstructed jet pseudo rapidity. The correction factors are derived and applied on jets that are corrected for pile-up offset.

- **Residual corrections for data:** The residual data to simulation scale factors are determined to account for the remaining small differences (of the order of 1%) between the jet response in data and simulation. It is determined after correcting the jets for pile-up offset and simulated particle response and applied in data only.

Jet Energy Resolution

The jet energy resolution also needs to be calibrated. Measurements show that the jet energy resolution (JER) in data is deteriorated, with respect to the simulated one. Therefore, the jets in MC need to be smeared to describe the data. The jet p_T resolutions are determined with dijet and photon+jet events. The reference resolutions obtained from simulation are parameterised as a function of particle-level jet p_T and average number of pile-up interactions in bins of jet η. Corrections for differences between data and MC simulation are applied as η-binned scale factors on MC.

Jet Flavours

Identification of jets emerging from a light flavor or from a heavy flavour quark is essential for many physics analyses. This plays a central role for having purer heavy-flavour signals and reducing the enormous backgrounds from light flavour processes. Jet flavours can be categorised as light flavour: originating from u, d, s quarks or gluons, charm: from c quarks, heavy flavour originating from b quarks. There are algorithms to identify jets originating from b quarks and also a new algorithm for c quark has been developed. In the following section b jet identification algorithm, which is used for the selection of b jets in this thesis, will be briefly explained.

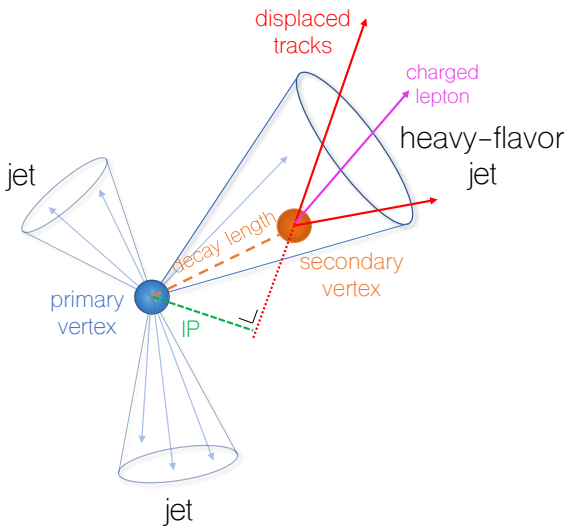

Figure 5.3: Illustration of a secondary vertex from the decay of the b or c hadron resulting in charged particle tracks (including possibly a soft lepton) that are displaced with respect to the primary interaction vertex, and hence with a large impact parameter (IP) value [105].

B-jet Identification

B-jet identification, named as b-tagging, relies on the properties of the production, (weak) decay of b-hadrons and also on the reconstructed PF elements. The relatively long life time of b-hadrons 1.5 ps ($c\tau \approx 450\mu$m) results in a distance that can be observed by high resolution tracking detectors. This displacement of a few millimetres with respect to the PV results in the presence of displaced tracks from which a secondary vertex may be reconstructed. This is illustrated in Fig. 5.3. Among the different algorithms developed and used by the CMS collaboration, Combined Secondary Vertex version 2 (CSVv2) [104] which is used as the b-tagger for the identification of b jets in this thesis, will be explained.

The tagging algorithm uses the information of PF jets, measured properties of all the charged particles in a jet, global muons and the PV. In addition to the track selection cuts, specific requirements are applied to have higher quality of tracks. These cuts are: at least 8 hits in the silicon tracker; the transverse momentum of the tracks to be greater than 0.8 GeV and the longitudinal component of the impact parameter to be smaller than 0.3 cm. Among the selected tracks, displaced tracks are identified as seeds if the value of the impact parameter is at least 50 μm and the corresponding significance is at least 1.2. For the reconstruction of the secondary vertices, the Inclusive Vertex Finder (IVF) uses the collection of reconstructed tracks with additional cuts mentioned above in the event. These tracks seed to clusters of nearby tracks depending on their minimum distance and the angles be-

tween them. The clusters are fitted with adaptive vertex fitter and vertices with low flight distance significance are removed. At this stage some tracks may belong to several vertices therefore some iterative quality cuts are applied [99] and the vertex is refitted. Secondary vertices are only considered when they have a mass of less than 6.5 GeV and that is not compatible with the mass of the K^0_S hadron in a window of 50 MeV. Additionally, the angular distance ΔR between the jet axis and the secondary vertex flight direction is required to be smaller than 0.3.

The CSVv2 algorithm is based on the CSV algorithm used in Run 1 and described in Ref.[106]. It combines the information of displaced tracks with the secondary vertices reconstructed with the IVF algorithm. The displaced tracks should satisfy the selections mentioned earlier in this section as higher quality cuts and these tracks should have an angular distance with respect to the jet axis, ΔR smaller than 0.3. At least two tracks per jet are required and if the there is a combination of two tracks with the mass of K^0_S meson in a window of 30 MeV are rejected. The training of the algorithm is performed in three vertex categories:

- jets with at least one secondary vertex
- jets with a "pseudo-vertex" : at least two tracks not compatible with a K^0_S meson and impact parameter significance[2] larger than 2
- no reconstructed secondary vertex or pseudo vertex associated to a jet

A neural network is used to combine the discriminating observables in each vertex category. These three categories are used to construct a likelihood ratio taking into account the fraction of jets of each flavour expected in $t\bar{t}$ events. The distribution of the discriminator values for the light, c and b jets is shown in Fig. 5.4. This distribution is used for defining the working points for b-tagging.

5.2.5 Missing Transverse Momentum

The total transverse momentum of the particles that do not interact with the detector material is obtained by balancing the momentum conservation in the transverse plane. It is defined as the negative vectorial sum over the transverse momenta of all PF particles and transverse energy is the associated scalar momenta of the PF particles.

Since there are jets involved for the missing transverse momentum calculation, missing energy also need to be corrected for jet energy scale. The jet-energy corrected missing transverse momentum is:

[2]Impact parameter significance is defined as the value of the impact parameter divided by its resolution, $\mathrm{IP}/\sigma(IP)$

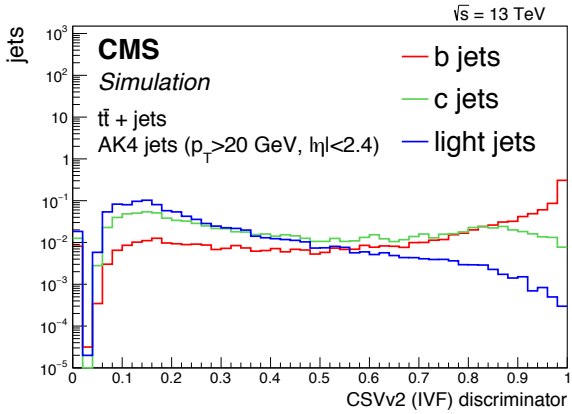

Figure 5.4: Distribution of the CSVv2 discriminator values for different jet flavours using $t\bar{t}$ events [105].

$$\vec{p}_{\mathrm{T,PF}}^{\,\mathrm{miss}} = -\sum_{i=1}^{N_{\mathrm{particles}}} \vec{p}_{\mathrm{T},i}(raw) - \sum_{j=1}^{N_{\mathrm{PFjets}}} (p_{\mathrm{T},j}^{\mathrm{corr}} - p_{\mathrm{T},jet}) \tag{5.5}$$

where the superscript "cor" refers to the corrected and "raw" to the uncorrected values.

Chapter 6

Event Selection

6.1 Introduction

In this Chapter, the selection of collision events enriched in $t\bar{t}Z$ with three leptons in the final state, hereafter referred as "ttZ event selection", candidates will be described. The event selection targets final states where one top quark decays to a W boson and a b-tagged jet and the W decays hadronically to two jets, while the other (anti)top quark decays to a W boson and a b-tagged jet and the W decays to a lepton and a neutrino and the Z boson decays to two charged leptons. This decay chain is shown in Fig. 6.1. This final state has the highest signal sensitivity with respect to the other $t\bar{t}Z$ decay channels, that are introduced in Sec. 2.3.2.

First identification and selections of the final state particles that are most

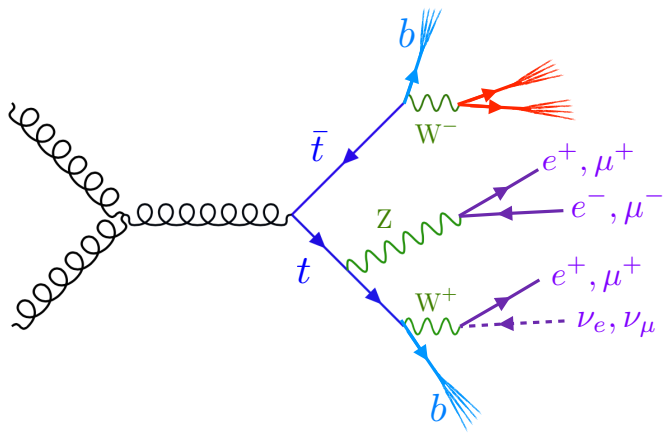

Figure 6.1: $t\bar{t}Z$ decay in three lepton final state. The jets are denoted by red, b-tagged jets by blue.

Table 6.1: MC samples and corresponding cross sections used, $t\bar{t}Z$ is the signal sample and all the other samples are used for background predictions. The NLO samples and Tune CUETP8M1 [107] are used, otherwise it is mentioned as LO.

Sample Name	Generator	σ (pb)
$t\bar{t}Z/\gamma^*$ To 2 Leptons, 2 Neutrinos	MadGraph-Pythia8	0.2728
$t\bar{t}W$ Jets To 1 Lepton, 1 Neutrino	MadGraph-Pythia8	0.2043
WZ To 3 Leptons, 1 Neutrino	MadGraph-Pythia8	4.67
$t\bar{t}H$ To not $b\bar{b}$ Quarks	MadGraph-Pythia8	0.2151
tZq To 2 Leptons	MadGraph-Pythia8	0.09418
$t\bar{t}\gamma$	MadGraph-Pythia8	3.697
tWZ To 2 leptons	LO MadGraph-Pythia8	0.01123
$t\bar{t}t\bar{t}$	MadGraph-Pythia8	0.009103
ZZ To 4Leptons	Powheg-Pythia8	1.256
WZZ TuneCUETP8M1	MadGraph-Pythia8	0.05565
WWZ	MadGraph-Pythia8	0.1651
ZZZ	MadGraph-Pythia8	0.01398
$W\gamma$ To 1 Lepton, 1Neutrino	MadGraph-Pythia8	585.8
$Z\gamma$ To 2Lepton	MadGraph-Pythia8	131.3

specific for $t\bar{t}Z$ production will be described. Then, the $t\bar{t}Z$ event selection will be discussed.

6.2 Data and MC Samples

In this analysis the data taken during 2016 that corresponds to an integrated luminosity of 36 fb^{-1} is used. The MC samples used for the signal and background estimations are normalised to the expected number of events corresponding to the total integrated luminosity, by Eq. 3.2. The cross sections used for the normalisations of the MC samples are calculated from the matrix element of the MC generator used. The samples and the corresponding NLO cross sections used in this analysis are listed in Tab. 6.2.

6.3 Physics Objets Identification

6.3.1 Lepton Selections

In $t\bar{t}Z$ events, charged leptons[1] come from the decay of a Z or W boson. These are called "prompt" leptons while those coming from the decays of light or heavy hadrons are referred to as "non-prompt" leptons. Leptons arising from instrumental effects causing misidentification of jets are called "fake" leptons. In this analysis the selection of leptons will be categorised as *tight* leptons that targets selecting prompt leptons and *loose* leptons that targets selecting non-prompt and fake leptons. These definitions will be used and explained in detail in Sec. 6.5.1.

Electron Selection

Electron p_T is required to be greater then 10 GeV and $|\eta|$ should be smaller than 2.5. The electrons (in fact leptons in general) in $t\bar{t}Z$ events carry relatively high momentum (see Fig. 6.2 for the p_{T} distribution of each lepton), thus there is almost no signal lost with this threshold. The pseudorapidity of 2.5 is constrained by the geometry of the detector. Tight electrons are selected using the electron MVA-based discriminator identification where the input variables are explained in Sec. 5.2.2. In this analysis the MVA discriminator threshold that corresponds to a electron selection efficiency of 90% is used. Applying the MVA discriminator is not sufficient to reject non-prompt electrons which have significant energy flow near their trajectories. Requiring electrons to be isolated from such nearby activity significantly reduces this background. The PF isolation is defined as the p_{T} sum over the charged hadrons, neutral hadrons and photons within a chosen $\Delta R = 0.3$ cone around the electron direction. In addition, the extra energy from pileup interactions present in the isolation cone is removed. Then the relative isolation is defined as the ratio of the isolation and electron momentum:

$$rel_{\mathrm{ISO}} = (\sum p_{\mathrm{T}}^{\text{charged had.}} + \max[0, \sum p_{\mathrm{T}}^{\text{neutral had.}} + \sum p_{\mathrm{T}}^{\gamma} - p_{\mathrm{T}}^{\mathrm{PU}}])/p_{\mathrm{T}}, \quad (6.1)$$

$$p_{\mathrm{T}}^{\mathrm{PU}} = \rho \mathrm{A}_{\mathrm{eff}} \quad (6.2)$$

where $p_{\mathrm{T}}^{\mathrm{PU}}$ corresponds to pile-up corrections, ρ is the event-specific average pile-up energy density per unit area and A_{eff} is the effective area which is the geometric area of the isolation cone scaled by a correction factor that accounts for the residual pseudo rapidity dependence on pileup. In order to reject secondary electrons produced in the conversions of photons in the

[1] In this chapter for simplicity leptons will be used as a shorthand for charged leptons

Table 6.2: Summary of the cuts applied on electrons. "Tight" column refers to the criterion applied for signal selection while the "Loose" refers to the selection of electrons for background prediction.

Cut	Tight	Loose
$\|\eta\| < 2.5$	✓	✓
$pt > 10$	✓	✓
$\|d_{xy}\| < 0.05$ (cm)	✓	✓
$\|d_z\| < 0.1$ (cm)	✓	✓
Id	MVA 90% eff.	MVA 95% eff.
relIso	$(\text{Barrel} : 0.0994, \text{EC} : 0.107) <$	< 1
$SIP_{3D} < 4$	✓	✓

tracker or electrons coming from hadrons there are additional cuts applied. These are cuts on the impact parameters[2] on the transverse d_{xy} and longitudinal plane d_z and on the significance of the three dimensional impact parameter (SIP_{3D}) which is the ratio of the three dimensional impact parameter and its uncertainty. Loose electrons are selected with applying looser cuts. All the cuts applied for tight and loose electron selection are listed in Tab. 6.2.

Muon Selection

Muons are required to lie in $|\eta|$ range of 2.4 and have p_T greater then 10 GeV. Prompt muons are required to pass the *medium* identification selections, which has more than 95% selection efficiency. These cuts are applied on PF muons and are listed below:

- To be reconstructed as a global or tracker muon: standalone muon tracks that are only reconstructed in the muon system are rejected
- Fraction of the valid tracker hits > 0.49 : track quality cut to avoid tracks with lost hits in the tracker
- Segment compatibility > 0.303 if muon is a Good Global Muon and > 0.451 otherwise: to assure the associated segments of the track match the expectation for a real muon

where a Good Global Muon passes the requirements below:

- To be reconstructed as a global muon
- Normalised global muon track $\chi^2 < 3$: To suppress hadronic punch-through and muons from decays in flight
- χ^2 for the standalone-tracker matching of local position < 12: quality cuts

[2]Distance to the vertex at the point of closest approach

- Value of the kink algorithm[3] applied to the global track kink finder normalised $\chi^2 < 20$: to further suppress muons from decays in flight

As in case of the electrons, isolation and impact parameter cuts are additionally applied for muon selections. The isolation definition is the same as in Eq.6.1, but calculated with a different cone size $\Delta R = 0.4$ and, a different pile-up correction, called DeltaBeta $\Delta\beta$, is used. Here the pile-up correction is made by summing the transverse momenta of the charged particles in the cone of interest coming from the pile-up. The PU correction term in Eq. 6.1 is $p_{\mathrm{T}}{}^{\mathrm{PU}} = 0.5 \sum p_T(PU)$ for muons. The factor 0.5 has been measured in jets [108] and empirically found that the total energy coming from neutral hadrons and photons is on average one half of the energy of the charged hadrons originating from PU. The cuts applied for tight and loose muons are listed in Tab. 6.3.

Table 6.3: Summary of the cuts applied on muons. "Tight" column refers to the criterion applied for signal selection while the "Loose" refers to the selection of electrons for background prediction.

Cut	Tight	Loose
$\|\eta\| < 2.5$	✓	✓
$pt > 10$	✓	✓
$\|d_{xy}\| < 0.05$ (cm)	✓	✓
$\|d_z\| < 0.1$ (cm)	✓	✓
$SIP_{3D} < 4$	✓	✓
PU corrected relIso	< 0.25	< 1
Muon Medium Id	✓	✓

6.3.2 Jet and b-tagged Jet Selection

PF jets with transverse momenta greater than 30 GeV and with pseudo rapidity $|\eta| < 2.4$ are selected. The cuts listed in Tab. 6.4 are used to remove fake jets arising from noise and spurious energy depositions in a single subdetector. As explained in Section 5.2.4 jets are corrected for contributions from PU and detector response (L1, L2, L3 corrections).

At PF jet reconstruction level, the final identification criteria of leptons are not yet applied. Therefore the reconstructed tracks or energy deposits in the calorimeters originating from leptons could be also used for jet clustering. Therefore, at analysis level this ambiguity need to be taken care by removing those jets that are actually formed by leptons. This is done in a cone of ΔR

[3]Kink algorithm looks for "kicks" along the reconstructed muon track that are caused when the muon interacts with the tracker. The track is evaluated from inward and outward directions, the χ^2 of the difference between the two states is evaluated.

of 0.4 around the jet. If a tight lepton, as identified in the previous sections, is found in this cone around the jet, the jet is removed from the jet collection. In order to select b-tagged jets medium working point determined by the CSVv2 algorithm is used. Jets having a CSVv2 value greater then 0.85 are defined as b-tagged jets. The b-tagging selection criteria used has a mistag rate of the order of 1% and a corresponding tagging efficiency of approximately 70% depending on the jet transverse momentum and pseudorapidity.

Table 6.4: Summary of the cuts applied for jet selection

Cut	Value
Neutral hadron energy fraction	< 0.99
Neutral electromagnetic energy fraction	< 0.99
Charged electromagnetic energy fraction	< 0.99
Charged hadron energy fraction	> 0
Charged multiplicity	> 0
Number of constituents jets	> 1
$\|\eta\|$	$<$2.4
p_T	> 30 GeV

6.3.3 Trigger Selection

For this analysis the trigger selection needs to be able to select events composed of three leptons and multi-jets originating from $t\bar{t}Z$. For this purpose, for example triggering on the transverse momentum of the jets is not a very good choice since it does not guarantee selecting events with leptons. One can use single lepton and double lepton triggers. In the case of double lepton triggers, in the HLT menu tighter isolation requirements are applied that may bias the non-prompt background estimation. Therefore only single lepton triggers are used for the selection of events. This choice is also motivated by the high p_T spectrum of the leading and subleading leptons that guarantees the trigger to be fired. Figure 6.2 shows the p_T spectrum of the three leptons selected by the analysis cuts (see Sec. 6.4) in a region dominated by the signal. The highest p_T lepton spectrum starts around 40 GeV that is high enough to fire the single lepton triggers with p_T thresholds of 27 and 24 GeV for electrons and muons, respectively.

The triggers used in data and simulation are shown in Tab. 6.3.3. As the nam-

Table 6.5: Single lepton triggers as they are defined in MC samples and in data.

	Data and MC
Single Electron	HLT_Ele27_WPTight_Gsf
Single Muon	HLT_IsoMu24 or HLT_IsoTkMu24

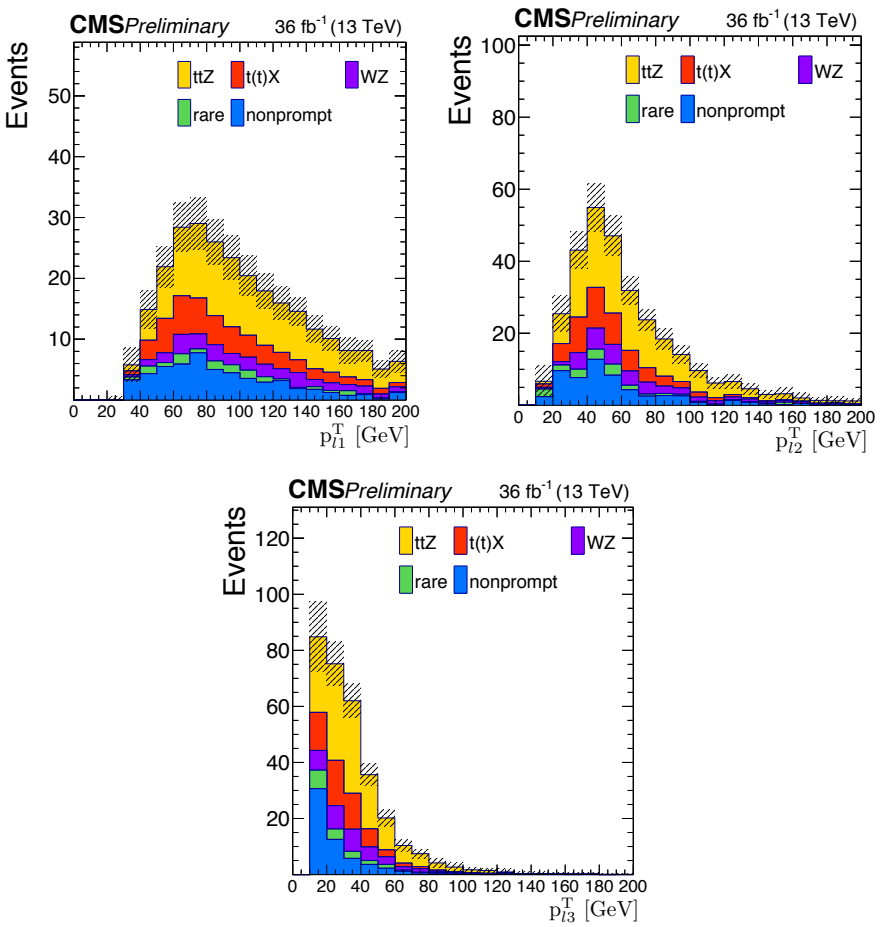

Figure 6.2: Transverse momentum spectrum of the leading (left), sub-leading(middle) and the trailing(trailing) lepton in a signal dominated selection region: Njets$\geq$ 4, $N_{\text{b jets}} \geq 2$. The yellow histogram corresponds to the $t\bar{t}Z$ signal and the other colours correspond to the backgrounds that will be explained in Sec. 6.5 .

ing convention of the triggers implies, a tight selection at trigger level which is looser then the electron identification cut, on electrons is applied and p_T of the electron is required to be greater than 27 GeV. In case of muons, at HLT level muons are required to be isolated or be an isolated tracker muon with p_T greater than 24 GeV. The corrections applied for trigger efficiencies are introduced in Sec. 6.6.1.

6.4 Event Selection

The $t\bar{t}$ pair in association with a Z boson is analysed in the three leptons final state. The corresponding decay chain is as follows:

$$pp \to t\bar{t}Z \to (t \to bW(\ell^{\pm}\nu))(t \to bW(jj))(Z \to \ell^{\pm}\ell^{\mp}), \tag{6.3}$$

In order to reconstruct $t\bar{t}Z$ production, the following selection of these events is performed:

1. Three prompt leptons passing the identification requirements described in Sec. 6.3.1 with p_T greater than 40, 20, 10 GeV are selected. The leading p_T threshold is decided to be 40 GeV to be highly efficient in trigger selection while, at the same time it does not eliminate signal events since the leading lepton p_T is above 40 GeV. The trailing lepton p_T is chosen to be 10 GeV to maximise signal events while not increasing the background that could decrease the sensitivity of the analysis.

2. Two of the leptons coming from the Z boson need to have the same flavour with opposite charges (SFOC) and form an invariant mass consistent with a Z boson in a 20 GeV mass window around the Z boson mass (81 and 101 GeV). In case of two SFOC pairs passing this selection the pair with the mass closest to the Z boson is selected.

3. At least two jets with the identification requirements described in Sec. 6.3.2 with p_T greater than 30 GeV.

Using the events that pass the selections of steps 1-3, various event categories are formed in different jet and b-tagged jet multiplicities: N_{jets} = 2, 3, and ≥ 4 with $N_{\text{b jets}}$ = 0, 1, and ≥ 2. In the final state it is expected to have at least four jets, two of which originate from b quarks. It is shown in $t\bar{t}Z$ simulation events that nearly 80% of the signal events fail the requirement of having all four jets with two of them being identified as b-jets. Therefore the lower jet and b-tagged jet multiplicity events are used to include a larger fraction of the signal events. Despite the larger background contamination, events with 3 jets, in particular the higher $N_{\text{b jets}}$ bins, improve the signal sensitivity, as this category recovers signal efficiency for the cases where one of the jets falls outside the acceptance. The N_{jets} = 2 category provides a background-dominated region that helps to constrain the background uncertainties. All the nine signal regions are used to extract the signal significance and the cross section. The jet multiplicity and b-tagged jet multiplicity distribution in $t\bar{t}Z$ events in MC are shown in Fig. 6.3.

6.5 Background Estimation

Processes with very similar or same final state particles and topology as the signal are called background events. One of the main goals of data analysis selection is to eliminate background events while keeping the signal acceptance at maximum possibly.

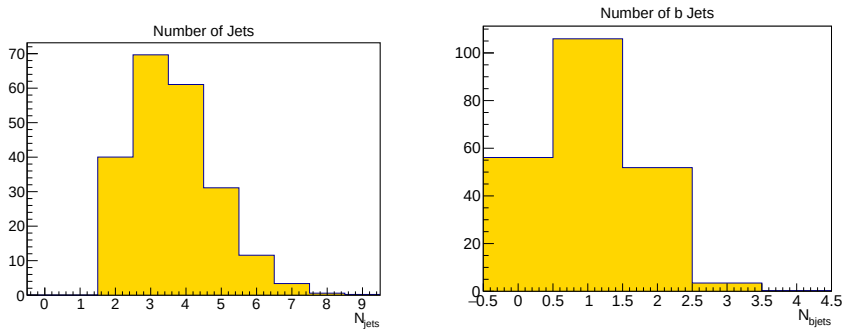

Figure 6.3: Jet and b-tagged jet multiplicity of the $t\bar{t}Z$ events from MC sample

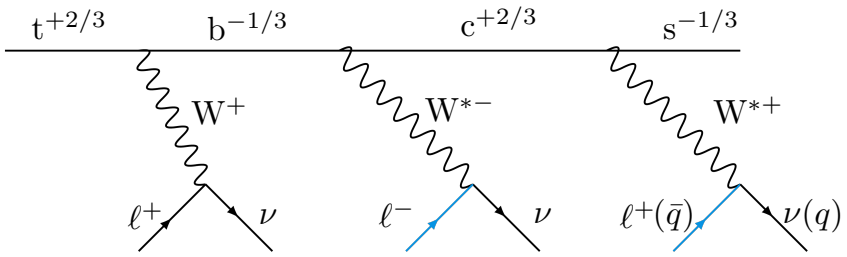

Figure 6.4: Semileptonic top quark decay, where the blue line refers to a non-prompt lepton.

In pp collisions there are other processes that have a similar final state as $t\bar{t}Z$. These are production of top-antitop quark pairs in association with bosons, multi-boson productions with three leptons and multi-jets in the final state. Depending on the number of non-prompt leptons present in the event we can categorise the backgrounds into the ones containing at least one non-prompt lepton and those where all three leptons are prompt. Backgrounds with three prompt leptons or four prompt leptons where one lepton escapes detection are estimated from MC simulation or from a control region defined in data. While backgrounds containing non-prompt leptons are estimated directly using data. These backgrounds will be explained in the following sections.

6.5.1 Backgrounds Containing Non-prompt Leptons

The determination of $t\bar{t}Z$ events targets selecting two SFOC leptons coming from the Z boson decay and the third lepton coming from the decay of a W boson. As defined before these leptons are called prompt leptons. They typically have large momentum ($p_T > 10$ GeV), are created at the primary interaction vertex and well isolated from other high momentum objects. Even there are identification criteria applied for selecting prompt leptons, there are some fraction of non-prompt leptons passing these identification crite-

ria. These non-prompt leptons may have a different source or they can be misidentified jets.

This background is measured in data because non-prompt decays and misidentification of jets may not be accurately modelled in MC simulations. For instance, when a top quark semileptonically decay, the non-prompt lepton may be coming from the decay of a virtual W boson which is originating from c or s quarks, as shown in Fig. 6.4. The modelling of such a processes may not be very accurate, due to the details of jet fragmentation and the underlying event models. Moreover, the probability of a jet to be identified as a lepton is very low therefore it is technically challenging to produce sufficient simulation events for the various distributions to be statistically reliable.

The sources of non-prompt leptons are different depending on the lepton type. For electrons, charged tracks in the trackers linked with clusters in the ECAL emerging from a jet or π^0 mesons can be reconstructed as an electron. This jet misidentified as an electron is called a fake lepton. Other sources of non-prompt electrons are photon conversions where a photon interacts with the detector material and converts to an electron positron pair, and leptonic heavy flavour decays where for example a bottom quark decays to a charm quark and a W^+ boson where W^+ boson decays to a positron and a neutrino. For muons the probability of misidentification of a jet as muon is almost zero due to the fact that jets can not reach the muon system which largely excludes the probability of a fake muon. The sources of non-prompt muons are heavy flavour decays (as in case of electrons) and mesons decay in flight i.e. where a Kaon meson decays to muon and neutrino.

The production of non-prompt leptons mainly occurs inside a hadronic jet where the lepton has nearby particles and as a consequence is not isolated. Therefore the main variable to differentiate non-prompt leptons from prompt leptons is the isolation. The leptons coming from a W boson or Z boson are typically isolated, whereas in a sample enriched in light and heavy flavour quarks, in a QCD sample, leptons are not isolated. This can be seen in Fig. 6.5 where the isolation distribution of muons and electrons are shown in a QCD sample where leptons are not isolated and in W+jets and DY MC samples where leptons are isolated.

The contribution of non-prompt lepton background to 3-lepton final state events are:

- Drell-Yann events where the two leptons from a Z decay are prompt and the third lepton is non-prompt;
- $t\bar{t}$ events where W bosons from both top and anti-top quarks decay into two prompt leptons and the third lepton is non-prompt or a minor part of events where $t\bar{t}$ decays semi-leptonically to one prompt lepton and the second and the third leptons are non-prompt.

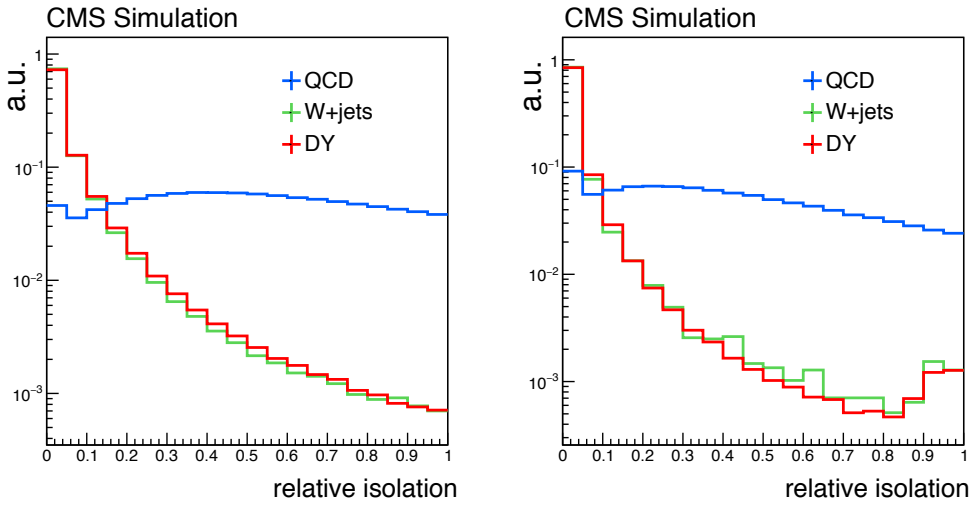

Figure 6.5: Isolation distribution of muons(left) and electrons (right) in QCD enriched, W+jets and DY MC samples.

The estimation of the non-prompt background is explained in detail in the following section. For fake of simplicity in the next sections and in the presentation of the results both the non-prompt lepton and fake leptons will be called as non-prompt lepton.

Fake Rate Measurement

In order to extract the fraction of non-prompt leptons that pass the identification requirements target at selecting prompt leptons, fake rate method is used. Fake rate is defined as the the probability for a loosely identified non-prompt lepton to pass the full set of tight requirements, designated as the tight to loose ratio:

$$FR = \frac{N_{tight}}{N_{loose}}, \tag{6.4}$$

where FR refers to fake rate, and $N_{tight,loose}$ to number of tight and loose leptons respectively. Loose leptons are selected a relaxed isolation requirement, and in case of electrons a looser identification with higher electron efficiency is applied. The loose and tight definitions for electrons and muons are given in Section 6.3.1. The idea is to measure how often a loose lepton passes tight lepton identification criteria in a sample enriched in fakes, and then to use this information to extrapolate from a sample consisting of loose leptons to a sample composed of tight leptons.

Fake rate needs to be measured in a control region that emulates the fake rate in signal phase space and do not contain prompt leptons not to bias

the measurement. Additionally the control region needs to be pure in non-prompt leptons to have good statistics. This can be done in a sample of events dominated by multijet events in data, called as well QCD enriched. This control region is called as the FR measurement region. Such QCD enriched events are selected using pre-scaled single lepton triggers. These single lepton triggers used for the selection are pre-scaled due to the high rate of events which are not possible to process due to the limitations of the DAQ. Pre-scaled trigger means only a fraction of the events are recorded and this fraction depends on the trigger menu. For instance in case of triggering on events with a threshold of 100 GeV on the p_{T} of the hadrons the pre-scale will be very low (i.e. 1/1000 meaning 1 event would be recorded among 1000 events) while in case of a threshold of 300 GeV the corresponding pre-scale would be larger (i.e. 1/10)[4].

The events in this control region are selected by requiring exactly one loose lepton and at least one additional jet where jets overlapping with leptons are not counted. The contribution of prompt leptons from W+jets events are suppressed by missing transverse energy $\mathrm{E_T^{miss}}$ and the invariant mass calculated using p_{T} $^{\mathrm{miss}}$ and p_{T} of each lepton, $\mathrm{M_T}$, cuts. This sample of events will be dominated by di-jet events where one of the jets is misidentified as a lepton.

The control region for the measurement of the FR is selected by applying the following cuts:

- One loose lepton
- At least one jet
- $\mathrm{E_T^{miss}} < 20$ GeV and $\mathrm{M_T} < 20$ GeV to suppress W+jet events
- Lepton candidates are required to be well separated with the leading jet

Despite the requirements on $\mathrm{E_T^{miss}}$ and $\mathrm{M_T}$ are applied to suppress prompt leptons in the measurement region, the contamination from electroweak processes remains, as can be seen in Fig. 6.6. This contamination of prompt leptons in the FR measurement region needs to be subtracted from the yields measured in data, since we would like to measure fake rate using non-prompt leptons. Due to the fact that the data events are selected with pre-scaled triggers, resulting less number of data events than expected, a scale factor is needed to be applied on MC simulation to agree with the pre-scaled data. This is done in the following way: in a control region orthogonal to the FR measurement region defined with $\mathrm{E_T^{miss}} > 20$ GeV and $70 < \mathrm{M_T} < 120$ GeV where the events are dominated by prompt leptons from W+jets events, a scale factor between data and MC simulation is obtained.

[4]These numbers are arbitrary chosen as an example

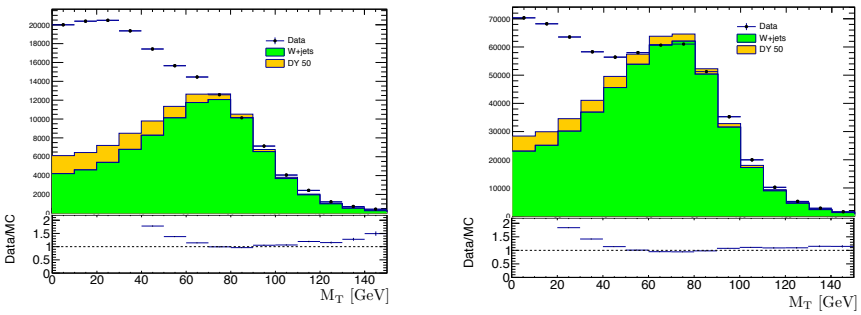

Figure 6.6: The normalised M_T distribution of the electrons (left), muons (right) and the missing energy. The prompt contribution of prompt leptons from W+jets and DY events are subtracted from the data in the measurement region, $M_T < 20$ GeV.

This scale factor is used to renormalise the prompt lepton contamination of DY and W+jets in the FR measurement region which are then subtracted. The corresponding plots are shown in Fig. 6.6.

The fake rate measurement need to be parametrised in terms of quantities sensitive to the event topology. In this way the extrapolation from the measurement to application region will be universal. Therefore fake rate is parametrised as functions of the η of the leptons and p_T^{cor}, where the latter is calculated to correct the lepton p_T as a function of the energy in the isolation cone:

$$p_T^{cor} = p_T \cdot (1 + max(0, iso. - iso_cut))$$

where Iso. is the isolation value of the loose lepton and Iso.Cut is the isolation cut applied for tight lepton selection that was described in Section 6.3.1. This definition [109] leaves the p_T of the leptons passing the isolation cut unchanged and modifies the p_T of those failing the tight isolation cut. This correction of the lepton p_T is used in order to fix the FR dependence on the flavor of the parton producing the non-prompt lepton, assuring that it is a better proxy of the corresponding parton p_T. As a result, the fake rate as a function of the mother parton p_T is flatter than the case where lepton p_T is used. The measured fake rates in data as a function of the corrected lepton p_T and η are shown in Fig. 6.7 separately for electrons and muons.

Implementation of the Fake Rate

Once the fake rate is measured, it is implemented to the events in the application region. The application region contains events that pass the full event selections with a different lepton composition. Among the three leptons, at least one of the leptons is required to pass the loose selection and to not to

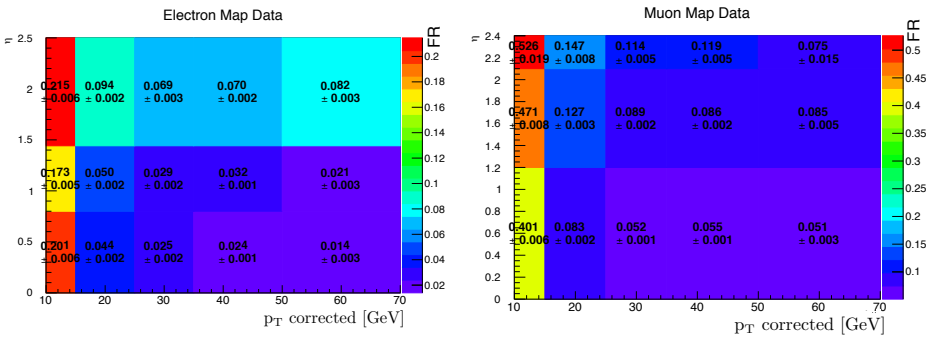

Figure 6.7: FR as a function of pt and eta for electrons(left) and muons(right) measured in data

pass the tight lepton criteria. These events are categorised as TTL, TLL, LLL where T,L refer to tight and loose but not tight leptons respectively. Each of these events are assigned a weight as a function of the $p_T(\text{cor})$ and η of the loose lepton to account for the probability of the non-prompt lepton to pass the prompt lepton requirements.

The TTL events are weighted by w_i:

$$w_i = \frac{FR_i(p_T, \eta)}{(1 - FR_i(p_T, \eta))}, \tag{6.5}$$

where i is an index denoting the loose but not tight lepton. The TLL events are weighted with $w_{i,j}$:

$$w_{i,j} = -\frac{FR_i(p_T, \eta)}{(1 - FR_i(p_T, \eta))} \times \frac{FR_j(p_T, \eta)}{(1 - FR_j(p_T, \eta))}, \tag{6.6}$$

where i and j are the indices denoting the two loose but not tight leptons. In order to extract the total number events containing of non-prompt leptons, the weights taken from the FR maps in Fig. 6.7 are summed over in all TTL and TLL events. The number of events with three leptons failing tight selection are negligible for this analysis therefore they are not considered. The event weights are calculating by extrapolating the number leptons from one base "loose-tight" to "non-prompt-prompt" as taking into account the efficiency of a loose (tight) lepton to be non-prompt (prompt) [110]. Furthermore, the efficiency of a prompt lepton to be a tight lepton need to be considered. It is shown from MC simulations that some fraction of the prompt leptons fail the tight lepton selection and fall into the loose lepton definition. These events are called spillage events and are predicted from MC simulation. They are subtracted from the total data driven estimated non-prompt events using the MC simulation. The total spillage events calculated in MC samples and the total data-driven background after the spillage are shown in Tab. 6.6, 6.7, 6.8.

	2 jets		
Process	0 bjet	1 bjet	$\geq$ 2 bjet
WZ	10.0 ± 0.4	0.7 ± 0.1	0.0 ± 0.0
ttX	0.8 ± 0.1	0.9 ± 0.1	0.1 ± 0.0
Rare	20.6 ± 1.8	1.9 ± 0.4	0.3 ± 0.2
ttZ	0.3 ± 0.1	0.5 ± 0.1	0.1 ± 0.0
total spillage	31.8 ± 1.8	4.0 ± 0.5	0.5 ± 0.2
Data-driven	197.9 ± 8.0	68.3 ± 4.9	3.2 ± 1.1
Data-driven after spil.	166.2 ± 8.2	64.4 ± 5.0	2.7 ± 1.1

Table 6.6: The total spillage, non-prompt background before and after the spillage subtraction for $N_{jets} = 2$ yields are shown, the errors are statistical only.

	3 jets		
Process	0 bjet	1 bjet	$\geq$ 2 bjet
WZ	3.3 ± 0.2	0.3 ± 0.1	0.0 ± 0.0
ttX	0.5 ± 0.1	0.7 ± 0.8	0.2 ± 0.0
Rare	7.8 ± 1.4	0.9 ± 0.3	0.2 ± 0.1
ttZ	0.4 ± 0.1	0.7 ± 0.1	0.3 ± 0.0
total spillage	12.0 ± 1.4	2.6 ± 0.4	0.7 ± 0.1
Data-driven	61.9 ± 4.8	19.7 ± 2.1	3.1 ± 0.8
Data-driven after spil.	50.0 ± 5.0	17.1 ± 2.1	2.4 ± 0.8

Table 6.7: The total spillage, non-prompt background before and after the spillage subtraction for $N_{jets} = 3$ yields are shown, the errors are statistical only.

	$\geq$ 4 jets		
Process	0 bjet	1 bjet	$\geq$ 2 bjet
WZ	0.64 ± 0.5	0.1 ± 0.1	0.0 ± 0.0
ttX	0.3 ± 0.0	0.4 ± 0.1	0.2 ± 0.0
Rare	1.3 ± 0.4	0.4 ± 0.2	0.1 ± 0.0
ttZ	0.6 ± 0.1	1.1 ± 0.3	0.7 ± 0.1
total spillage	2.8 ± 0.6	2.1 ± 0.4	1.0 ± 0.1
Non-prompt	17.5 ± 2.2	9.2 ± 1.4	3.5 ± 0.9
Data-driven after spil.	14.7 ± 2.3	7.1 ± 1.5	2.5 ± 0.9

Table 6.8: The total spillage, non-prompt background before and after the spillage subtraction for $N_{jets} \geq 4$ yields are shown, the errors are statistical only.

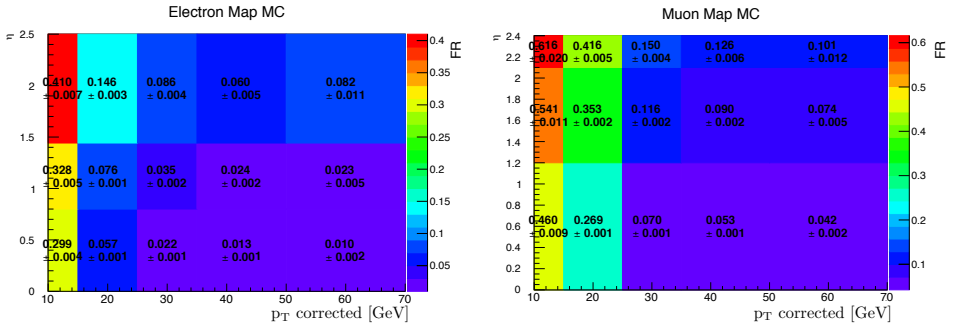

Figure 6.8: FR as a function of pt and eta for electrons(left) and muons(right) measured in QCD simulation

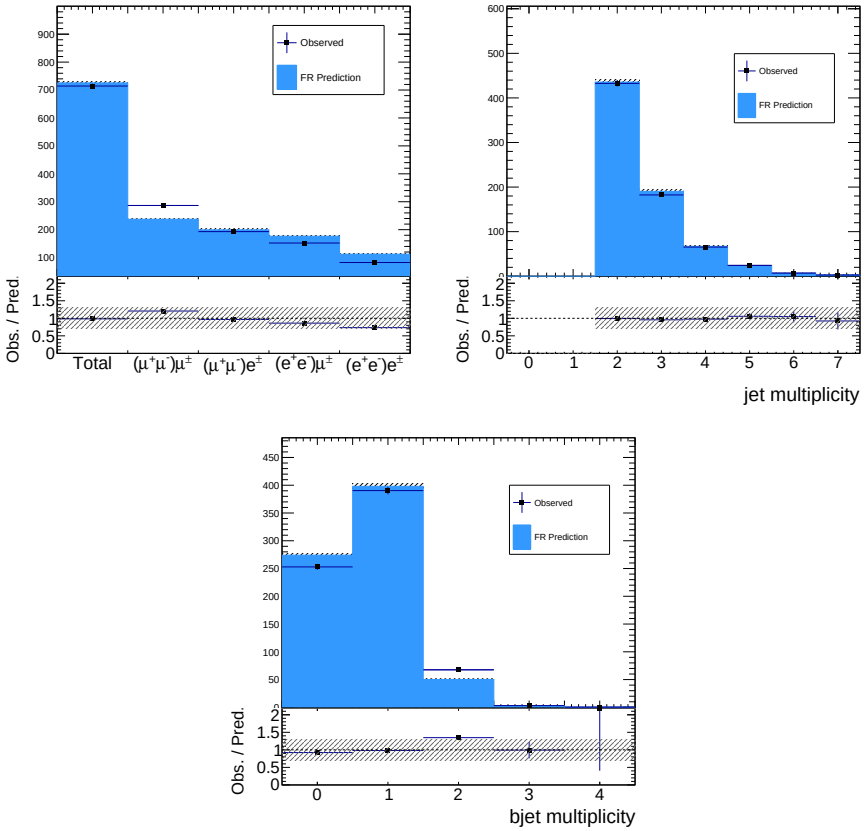

Figure 6.9: Fake background estimation closure with fake rate measured in QCD measurement region. Shown are the distributions of the predicted number of events as a function of the flavour composition of the event, (b-)jet multiplicity N_{jets} ($N_{\text{b jets}}$) for events with muons and electrons as obtained from a $t\bar{t}$ MG5_aMC@NLO sample. 30% uncertainty covers the discrepancy between observed and predicted yields (light hashed)

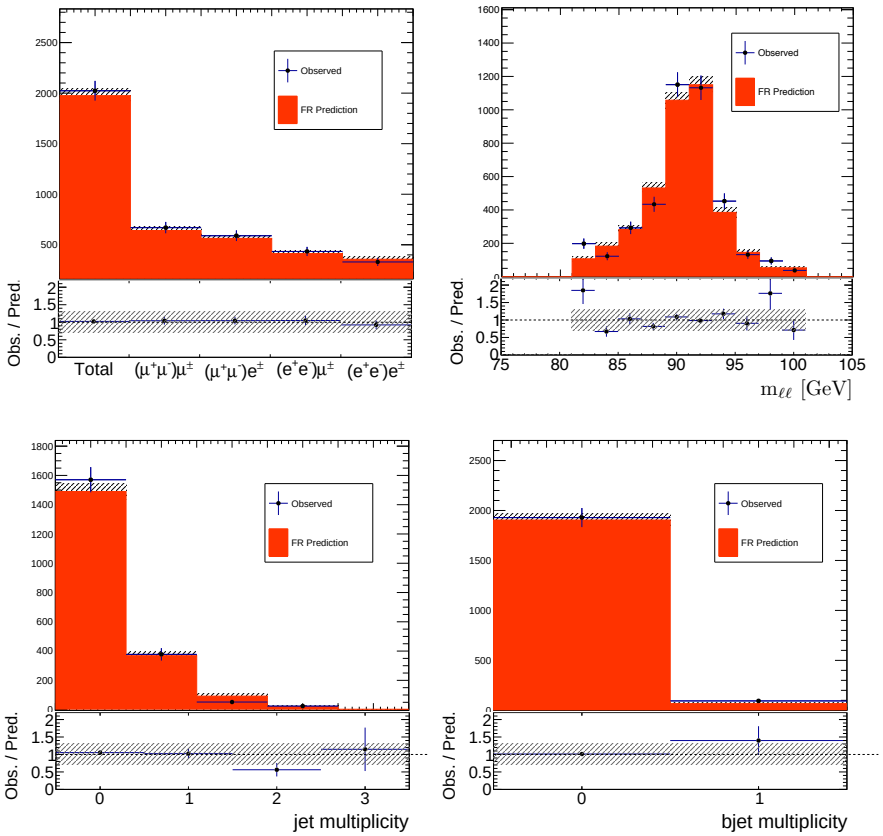

Figure 6.10: Fake background estimation closure with fake rate measured in QCD measurement region. Shown are the distributions of the predicted number of events as a function of the flavour composition of the event, invariant mass distribution of the SFOC lepton pair, (b-)jet multiplicity N_{jets} ($N_{\text{b jets}}$) for events with muons and electrons as obtained from a DY MG5_aMC@NLO sample. 30% uncertainty covers the discrepancy between observed and predicted yields (light hashed).

Validation of the Fake Fate Method

The fake rate method relies on the fact that the fake rate measured in QCD events can be transferred to $t\bar{t}$ and DY events in a different phase space. This can be verified using MC simulation and also data in a defined control region. This procedure is called as a closure test.

We perform a closure test using the simulated events. The distributions and yields of the events containing TTT leptons, called as observation in MC, are compared with ones where the fake rate weights are applied on events composed of TTL and TLL leptons, called as prediction. The degree of the discrepancy between the observation and prediction is reflected as the systematic uncertainty of the non-prompt background prediction. In order to be consistent, the fake rate is measured in QCD MC simulation with the

same selections applied as in data. The measured FRs in QCD simulation as a function of the p_{T} and η of the lepton are shown in Fig. 6.8

Due to the fact that non-prompt leptons in the 3 lepton final state have two different origins, $\mathrm{t\bar{t}}$ and DY, the closure tests are performed for both separately. The non-prompt contribution due to DY events are selected with 3 leptons where 2 of them are SFOC which have an invariant mass of $(m_Z \pm 10)$ GeV, while the contribution due to $\mathrm{t\bar{t}}$ events are selected with 3 leptons where 2 of them are SFOC and with the requirement of at least 2 jets. Figure 6.9 and 6.10 show the distributions of several event variables for the DY and $\mathrm{t\bar{t}}$ control regions: event yields in flavour composition, multiplicity of jets and b-tagged jets and the invariant mass distribution of the SFOC lepton candidates coming from the Z boson (only for the DY control region). The agreement between the FR predictions and MC observations are generally good while the discrepancies in some cases could be up to 30%. This will be taken as a systematics uncertainty and will be referred in detail in Section 7.2.

The fake background estimation method is also validated using data. The non-prompt contribution due to DY events are selected with the requirements applied on MC simulation. Additionally the cuts on $E_T^{miss} < 30$ GeV, $M_T < 30$ GeV are applied in order to enhance the DY events. As in case of MC, the same selections are applied to $\mathrm{t\bar{t}}$ events with the selection of events outside the invariant mass of the Z boson $m_{ll} < m_Z - 10, m_{ll} > m_Z + 10$ GeV with at least one b-tagged jet. Figure 6.11 and 6.12 show the distribution of several important event variables. The agreement between the data and prediction in DY dominated events is in quite good agreement. In $\mathrm{t\bar{t}}$ dominated events a discrepancy between the data and prediction is observed. This is related to the fact that the sources of these two non-prompt leptons are different and in case of $\mathrm{t\bar{t}}$ events the method does not work fully. Thus, the discrepancy is covered by the systematic uncertainty assigned to the method.

6.5.2 Backgrounds with only Prompt Leptons

Events with three lepton final states with jets occur due to multiboson productions or top quark (pairs) in association with a W, Z or Higgs boson. These processes are categorised as:

- WZ
- Top (pair) production with bosons ($\mathrm{t\bar{t}X}$): $\mathrm{t\bar{t}H}$, ttW, tWZ, tqZ, $\mathrm{t\bar{t}t\bar{t}}$
- Rare SM processes: ZZ, $\mathrm{Z}\gamma^*$, $\mathrm{W}\gamma^*$, WZZ, WWZ, ZZZ

WZ events have three leptons in the final state coming from the leptonic decay of a W boson to a lepton and a neutrino and a Z boson to two leptons

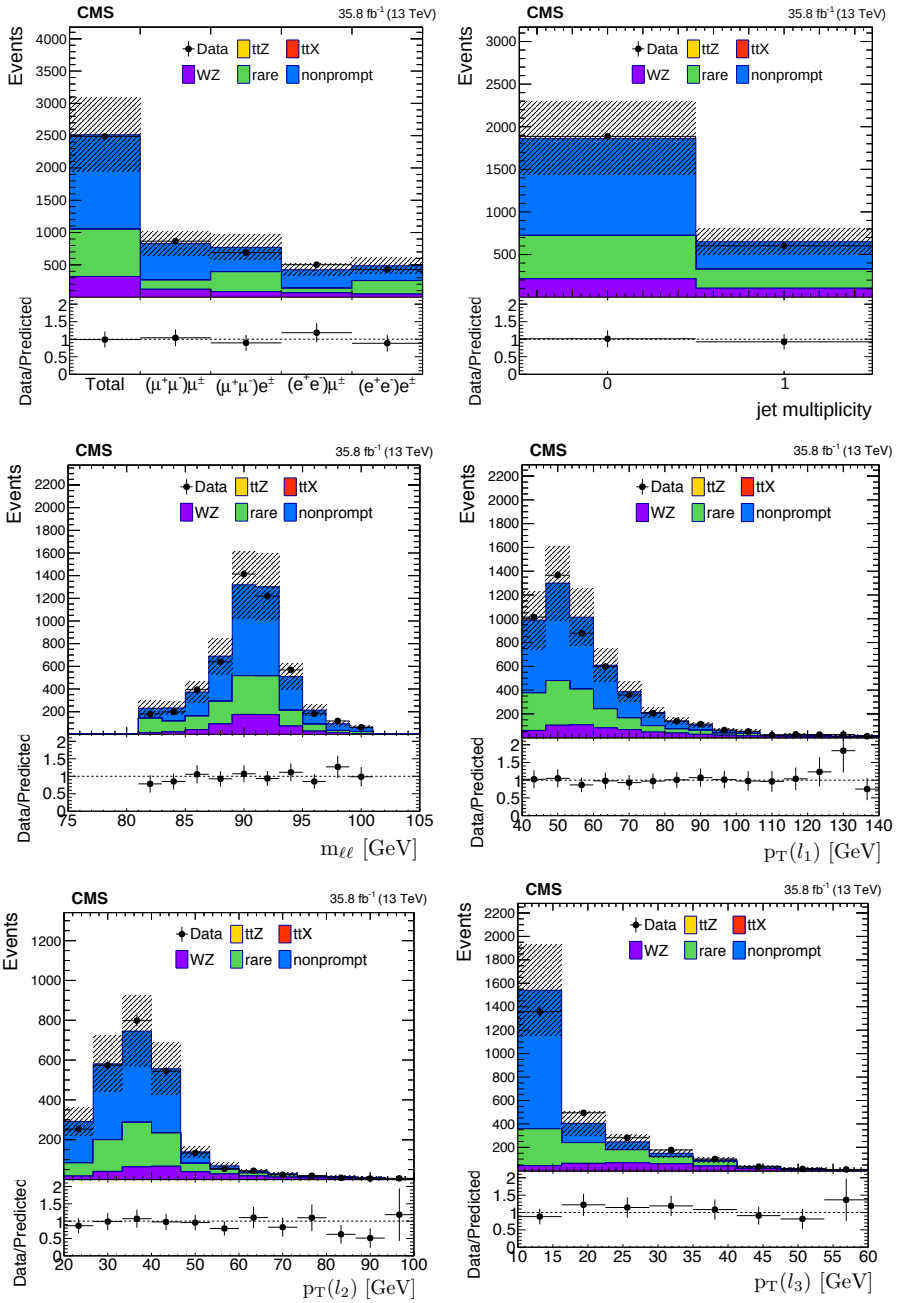

Figure 6.11: Fake background estimation closure with fake rate measured with data-driven method where the fakes are dominantly due to DY. Yields are normalised to an integrated luminosity of with 35.8 fb^{-}1. The shaded band represents the total uncertainty in the prediction of the background and the signal processes.

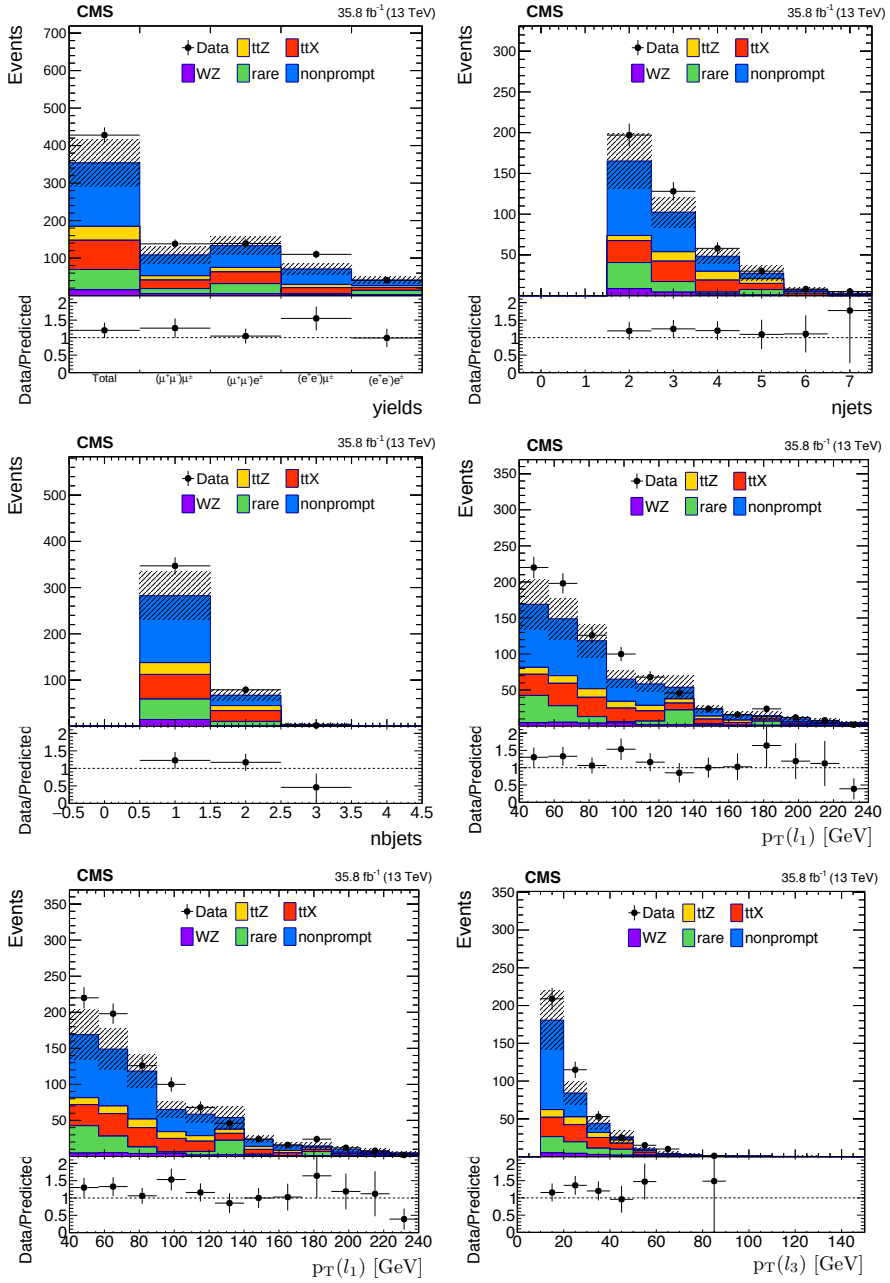

Figure 6.12: Fake background estimation closure with fake rate measured with data-driven method where the fakes are dominantly due to $t\bar{t}$. Yields are normalised to an integrated luminosity of with 35.8 fb^{-1}. The shaded band represents the total uncertainty in the prediction of the background and the signal processes.

with same flavour and opposite charge. This background has the highest expected number of events with 0 bjet and non-negligible contribution in events with b-tagged jets. The data used for this analysis contains a substantial number of WZ events that can be purely selected and compared with the MC predictions. Therefore a control region of pure WZ event selections are defined. In this control region the predicted yields and observed yields from data is compared with the simulation and if needed scale factors are applied to simulation. The WZ events are selected by applying the following requirements:

- Three leptons pass nominal identification and isolation cuts. p_T thresholds remain the same as in $t\bar{t}Z$ signal selection.
- Two leptons form a SFOC pair with $|m_{\ell\ell} - m_Z| < 10$ GeV in order to reconstruct lepton pairs decaying from a Z boson candidate
- $E_T^{miss} > 30$ GeV in order to select neutrino's missing energy from a W boson decay
- Transverse mass of third lepton and missing energy is required to be at least 50 GeV to suppress contamination from Drell-Yan process, $M_T^{\nu\ell} > 50$ GeV

With this selection a data sample 85% pure (WZ/WZ + *background*) in WZ is obtained. The expected background containing non-prompt leptons is measured from data using the method described in the previous chapter. The other background contributions are obtained from simulated samples. We observe an overall reasonable agreement between the data and the total expectation in all four lepton flavour channels and also in the kinematic distributions. The ratio of the observed to predicted yields is found to be 0.99 ± 0.07 where the uncertainty is only statistical. Fig. 6.13 shows the number of events in each lepton flavour channel and the N_{jets}, and $N_{\text{b jets}}$ distributions. With this level of agreement between data and the MC prediction, we proceed without applying a scale factor to the WZ prediction obtained by MC, and use this study to assess the systematic uncertainty in the WZ background prediction. In Fig. 6.14 the kinematical distributions $m_{\ell\ell}$ of the Z boson candidate, $M_T^{\nu\ell}$ transverse mass of third lepton and missing energy are shown.

The processes in $t\bar{t}X$ category contributes mainly to events with b-tagged jets due to the decay of top quarks to b-tagged jets. The tZq events with the following decay chain:

$$tqZ \rightarrow bW(\ell^{\pm}\nu)Z(\ell^{\pm}\ell^{\mp}), \tag{6.7}$$

and tWZ with the following decay chain:

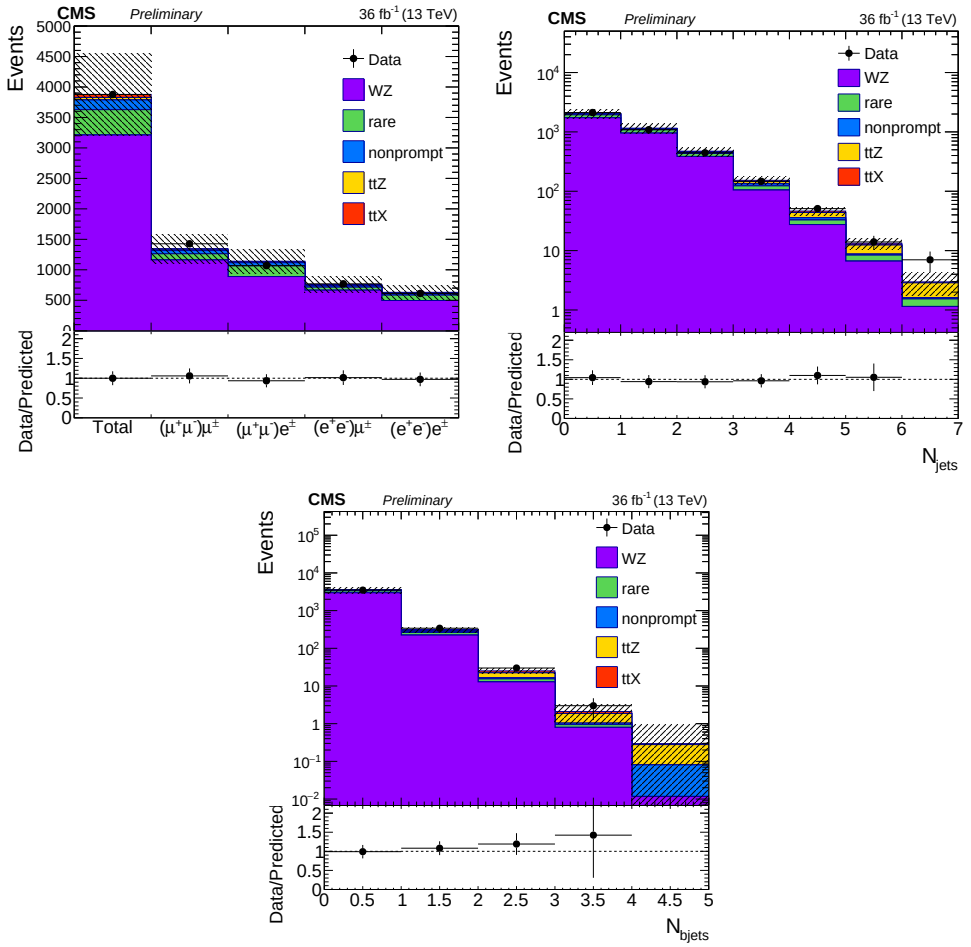

Figure 6.13: WZ control region plots: Distributions of the total yields versus lepton channel, jet and bjet multiplicity where the bottom plots show the data and prediction ratios. The shaded band represents the total uncertainty in the prediction of the background and the signal processes.

$$tWZ \rightarrow bW(\ell^{\pm}\nu)W(q\bar{q})Z(\ell^{\pm}\ell^{\mp}) \tag{6.8}$$

has a similiar topology to $t\bar{t}Z$ events. The difference comes from the fact that in $t\bar{t}Z$ events, it is expected to have two b-tagged jets. Considering the fact that the signal is exploited in 0, 1 and 2 and more b-tagged jets regions tqZ and tWZ events fall into the signal selections with the same topologies.

In case of $t\bar{t}H$ events where H boson decays to $Z^{*}Z$ has the same event topology as $t\bar{t}Z$ with additional jets. The events where H boson decays to $W^{*}W$ has three leptons with 2 b-tagged jets but has less contribution since there is no Z boson in the final state. These processes with very similar topology do not contribute to the event yields as high as $t\bar{t}Z$ since they have lower

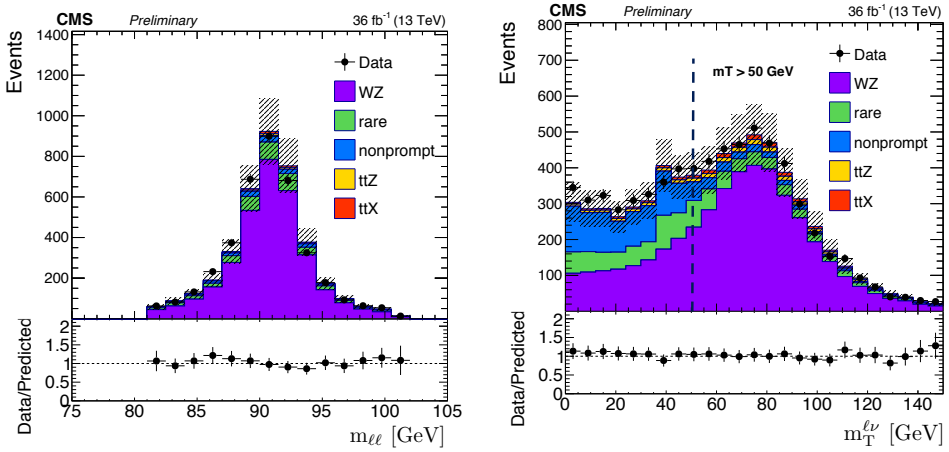

Figure 6.14: WZ control region plots: Distributions of the reconstructed invariant mass of the Z boson candidate and transverse mass of third lepton and missing energy where the bottom plots show the data and prediction ratios. The shaded band represents the total uncertainty in the prediction of the background and the signal processes.

cross section values at 13 TeV. The most significant background contribution in this category is from tZq which has a higher cross section with respect to tWZ and the latter has a higher cross section than $t\bar{t}H$. The corresponding cross sections are given in Tab. 6.2.
The other process in this category is the production of two top quark pairs that gives 3 leptons where 3 tops decay leptonically. But this background has a very small effect due to the smaller cross section and also due to the absence of Z boson in the final state.

Rarer processes, like multi-boson processes, contribute via fully leptonic decays of bosons. These can in principle give three and more leptons. These events pass the signal selection when one of the leptons either escape detection or fail the lepton selection. Their contribution to background events is higher in 0 b jet event category but also not negligible in events with b-tagged jets. The $t\bar{t}X$ and rare processes are estimated from simulation scaled by their NLO cross section and normalised to the integrated luminosity. The yields obtained from simulation are adjusted using scale factors to take into account several effects that lead to data and simulation differences. This is detailed in the next section.

6.6 Corrections Applied to Simulation Yields

The generated MC samples may not describe the data very well due to several factors. These are due to the imperfect modelling of the physics pro-

cesses and also due to differences between the conditions implemented for detector simulation and data-taking conditions. MC samples are generated with the implementation of the current state of the detector, but during data taking the detector configurations may change due to the run conditions. Additionally changes in the beam conditions, pile up and also the ageing of the detector need to be taken into account. Therefore this difference is corrected in simulation by applying scale factors (SF) and systematic uncertainties associated with these SFs are estimated. The relevant SFs for this analysis will be described in the following sections.

6.6.1 Corrections for Trigger Efficiencies

Triggers are applied both in data and in simulation, which can have different efficiencies due to different data taking conditions and also due to the different pile up and run conditions. Therefore this difference between data and simulation needs to be eliminated. This is done by measuring the trigger efficiencies in simulation and data and then extracting a correction factor. First, single lepton trigger efficiencies in W+jet simulation, where exactly one electron (muon) passes lepton requirements, are evaluated. Then the efficiency measurement in data is performed using a data sample collected with triggers that are not correlated with the single-lepton triggers. In this case we used data that is triggered with a MET requirement. In order to suppress non-prompt lepton contribution a cut of $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV and $\mathrm{M_T} > 20$ GeV is applied in data, as well as in simulation. Trigger efficiencies for single-lepton events differ nearly to 10% in data and simulation and shown in Fig. 6.15 as a function of the lepton p_T, η. It can be seen that the single lepton trigger efficiency measured in data and simulation do not agree very well in low p_{T} values while a better agreement is observed for p_{T} values greater than 40 GeV. Considering the efficiency measurement as a function of the lepton pseudorapidity, the agreement between the data and simulation is not ideal. This is due to the fact that the selected leptons are dominated by the low p_{T} ones causing a discrepancy. In case of the leptons coming from the $\mathrm{t\bar{t}Z}$ production, the highest p_{T} lepton spectrum starts around 40 GeV, Fig. 6.2. Therefore, the analysis is not affected by this discrepancy caused by the low p_{T} leptons.

The single lepton trigger efficiencies are shown in Fig. 6.16 in as a function of η and p_T. In order to extract three-lepton efficiency depending on the p_T and η of the selected three leptons we calculate a "weight" or an efficiency per event using the formula given below:

$$\text{Total Eff} = 1 - [(1 - \varepsilon_1(p_T, \eta)) \cdot (1 - \varepsilon_2(p_T, \eta)) \cdot (1 - \varepsilon_3(p_T, \eta))], \quad (6.9)$$

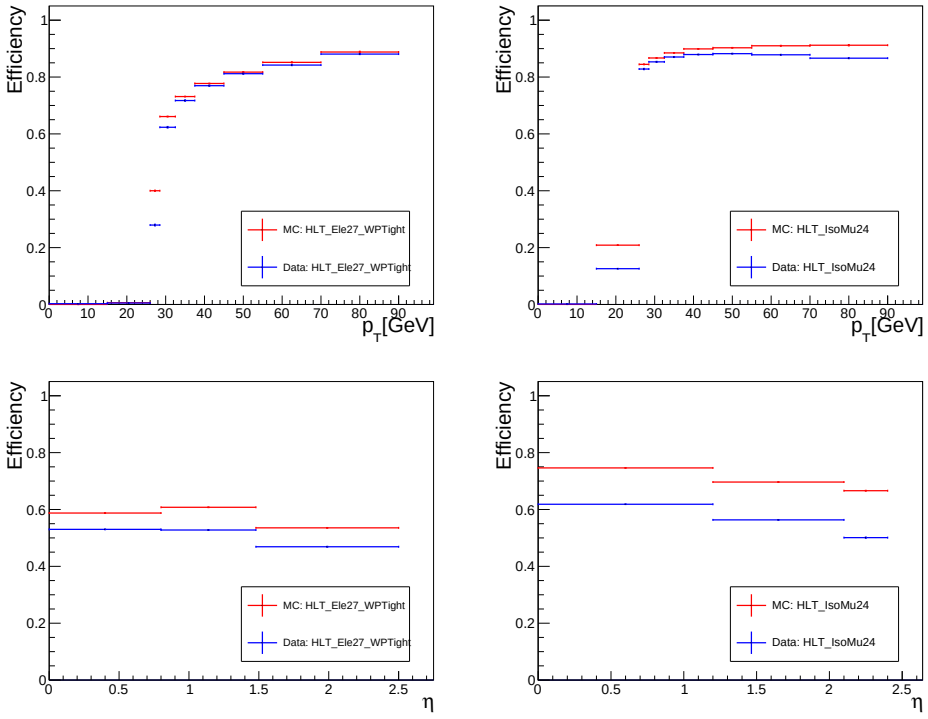

Figure 6.15: Measured efficiencies for single-electron(left) and single-muon (right) triggers as a function of lepton p_T, η.

where $\varepsilon_i(p_T, \eta)$ is taken from the efficiency maps in Fig. 6.16, 6.17 either from data or from simulation. The validation of the effective trigger efficiency for events with 3-leptons using single-lepton trigger efficiency using Eq. 6.9 is demonstrated in Fig. 6.18. Here the trigger efficiency for events with 3-leptons is measured using Eq. 6.9 where the trigger efficiency is taken from the efficiency maps as a function of the lepton p_T and η. The trigger SFs are the ratio of the total efficiency extracted using data "blue points" and total efficiency extracted using simulation (W+jet) red points shown in Fig. 6.18. In addition, the validation of the method is made directly measuring the efficiency by selecting 3-lepton events in WZ using simulation and shown by the green points in Fig. 6.18. As seen, the overall trigger efficiency for events with 3-lepton events is more than 97%. The SF to correct the small difference between data and simulation are applied to the MC samples.

6.6.2 Corrections for Lepton Selection Efficiencies

As in the case of trigger efficiencies the lepton selection efficiency may slightly differ between data and simulation. The efficiency of lepton selection is measured using a Tag-and-Probe method [111]. The Tag-and-Probe

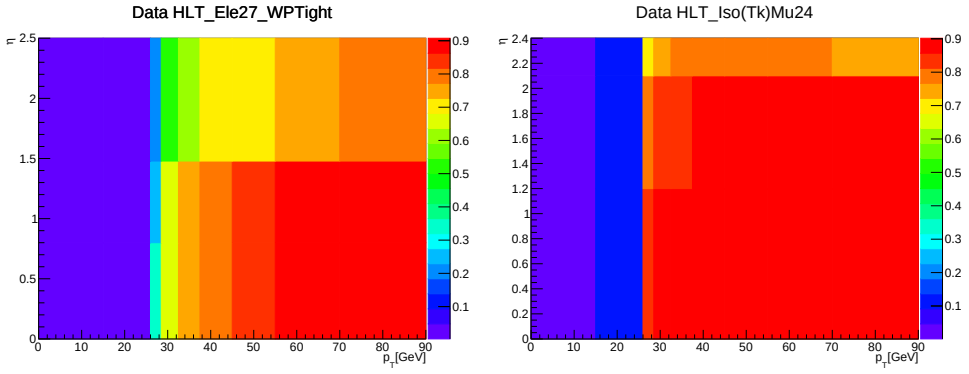

Figure 6.16: Single electron (left) and muon(right) trigger efficiency measured in data as a function of η and p_T

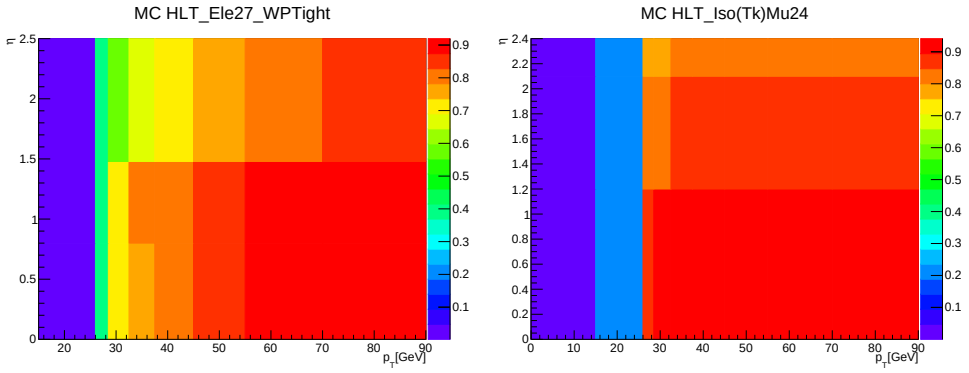

Figure 6.17: Single electron (left) and muon(right) trigger efficiency measured in simulation as a function of η and p_T

is a generic tool developed to measure any defined object efficiency by exploiting di-object resonances like Z or J/Psi. The Z boson is reconstructed within the invariant mass range of 60-120 GeV of the two objects of which one lepton passing a tight selection called the tag, and the other lepton passing a looser selection which is called the probe. The probe selection is defined according to the object that the efficiency to be measured. The tag and passing probe and all the possible lepton pairs (tag+passing probe and tag+failing probe) invariant mass shapes are fitted with the signal and background shapes. The efficiency is computed by the ratio of the number of signal lepton pairs that both pass tag and probe selection with the number of all lepton pairs. The same method is applied to both data and Z boson simulation. The SF is obtained by taking the ratio of the efficiency measured in data with measured in simulation and applied to simulation as a function of pseudorapidity and transverse momentum. The effect of the lepton SFs is shown in Fig. 6.19 where the left plots show the distribution without the SF applied and right with the SF applied.

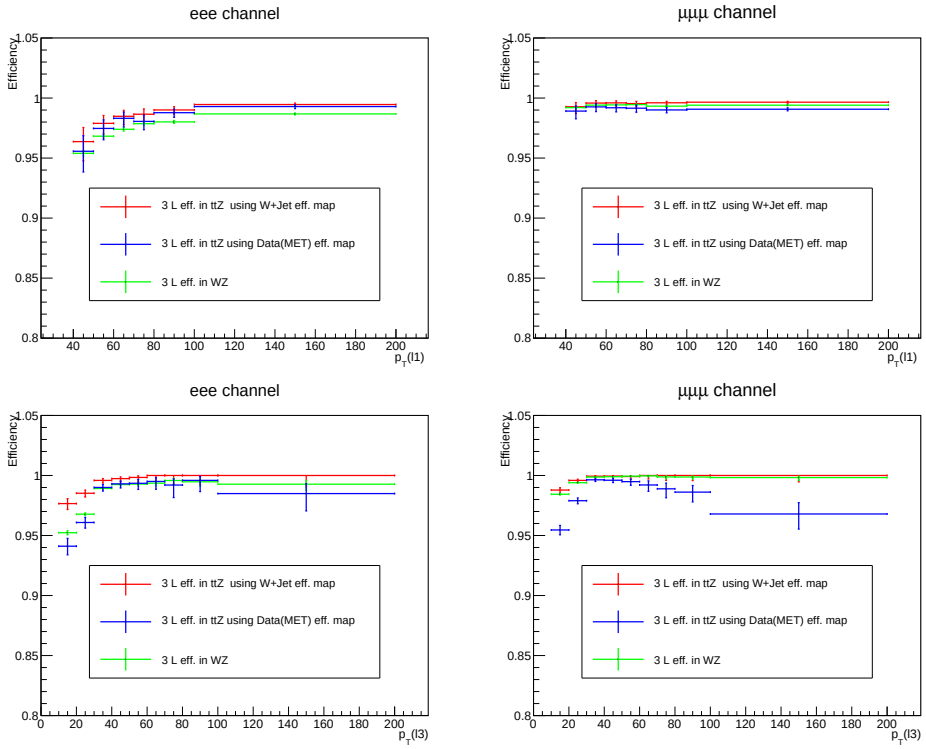

Figure 6.18: Trigger efficiencies for *eee* (top) and $\mu\mu\mu$-channels (bottom) as a function of the leading (left) and trailing lepton(right) p_T measured using the efficiency maps taken from data and simulation and also measured directly in WZ MC sample.

6.6.3 Corrections for Pile-up

The actual number of pile-up interactions in data is only known after the data taking while the pile up distribution in simulation is randomly sampled from a poisson distribution with a mean equal to the expected number of interactions in data. Therefore the simulation needs to be reweighed to correctly describe the PU in data. The distribution of the number of pile-up interactions in data and simulation before and after PU reweighing is shown in Fig. 6.20.

6.6.4 Corrections for b-tagging Efficiencies

The distribution of the btagging discriminator, CSVv2, is reweighted to correct the shape for light and heavy flavour jets in MC simulation to match the one measured in data.. The b-tagging efficiency scale factors for b and light flavour jets using the Tag and Probe method, explained in Sec. 6.6.2, using jets passing btag discriminator value as the tag and the second jets as the probe in $t\bar{t}$ dominated control region. In this control region, both W bosons coming from top quarks decay to leptons, precisely one to electron

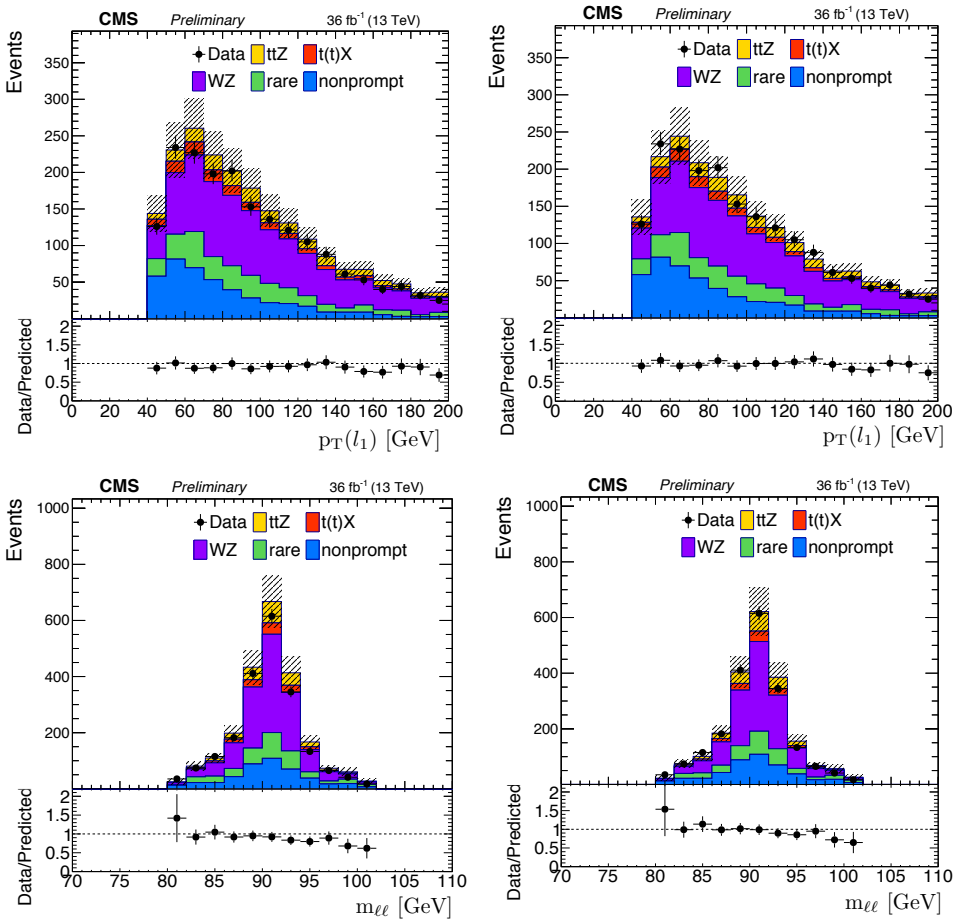

Figure 6.19: The distributions of the leading lepton transverse momentum and invariant mass of Z boson candidate left without lepton SF applied in MC simulation, right with lepton SF applied in MC simulation. The bottom plots show the ratio of data and prediction.

one to muon, and both of the jets coming from the top pairs are expected to be b-tagged jets. The SF depend on the jet flavour, transverse momentum and pseudorapidity and applied in the following way. The probability of a given configuration of jets in MC simulation and data is defined as:

$$P(MC) = \prod_{i=tagged} \varepsilon_i \prod_{j=not\ tagged} (1 - \varepsilon_j) \tag{6.10}$$

$$P(Data) = \prod_{i=tagged} \mathrm{SF}_i\varepsilon_i \prod_{j=not\ tagged} (1 - \mathrm{SF}_j\varepsilon_j) \tag{6.11}$$

where ε is the MC btagging efficiency and SF is the ratio of btagging efficiency in data and MC. The event weight is calculated as the ratio of these

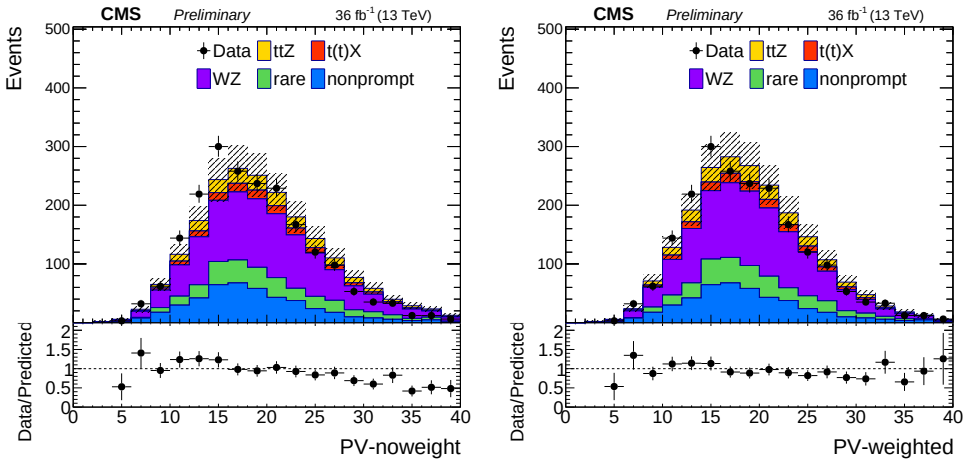

Figure 6.20: The distributions of the number of pile-up interactions left without PU correction in MC simulation, right after PU reweighing in MC simulation. The bottom plots show the ratio of data and prediction .

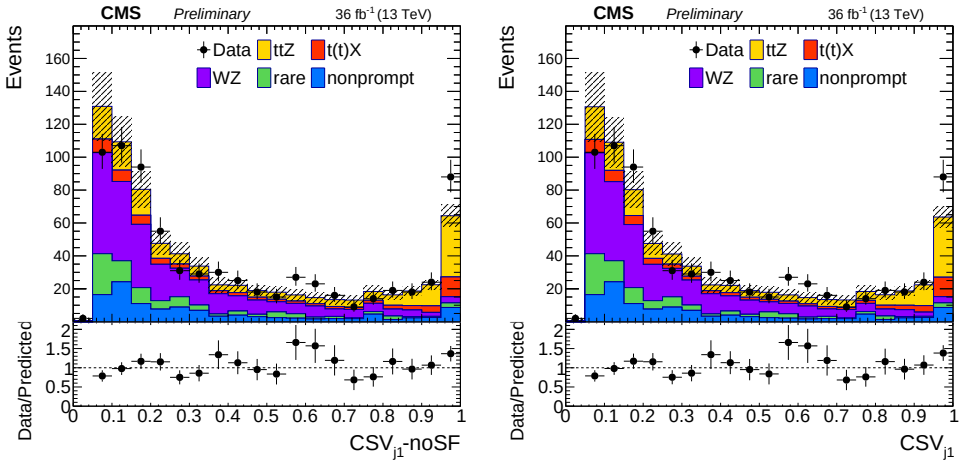

Figure 6.21: The distributions of the CSVv2 distribution of the leading jet left without btagging SF applied in MC simulation, right with btagging SF applied in MC simulation. The bottom plots show the ratio of data and prediction.

probabilities in data and simulation. The CSVv2 distribution of the leading jet with and without b-tagging SFs are shown in Fig. 6.21. Due to the fact that the SFs are very small that they do not make a visible effect when they are applied.

6.7 Event Yields and Distributions

Once all the backgrounds are defined and the corrections applied on simulations are described, we can evaluate the event kinematics and related

distributions. The event selections are applied to the signal events and background processes where the effect of each cut on the signal and background events are shown in Table 6.9. As can be seen in the table the backgrounds are eliminated up to 0.5% with the applied cuts. The contributions of each background source in each b-tagged jet multiplicity are shown in Table 6.10. In 0-bjet category the background is dominated by WZ events, in 1-bjet category by background processes with non-prompt leptons and in 2 and more b-tagged jet categories the dominant background is $t\bar{t}X$.

The distributions of the observed and predicted signal and background yields for the main analysis selection $N_{jets} \geq 2$ are shown in Fig. 6.22, while the distributions of the most sensitive analysis categories $N_{jets} \geq 3$ and $N_{bjets} \geq 1$ are shown in Fig. 6.23. Finally the total number of signal, background and data yields in all analysis categories obtained by the methods described in this chapter are given in Table 6.12-14.

The agreement of the yields and distributions in data and prediction agree quite well in most of the analysis categories. While a discrepancy is seen in $N_{\text{b jets}} = 2$ bin with 2 and 3 jets events. This discrepancy is investigated in detail for potential mis-modelling of the backgrounds, b-tagged jet efficiencies and also for any dependence on data taking run periods or on used the b-tagged algorithm. As a conclusion of these studies, no hint of any mis-modelling or any other dependency was found that this is attributed as a statistical fluctuation in data.

Cuts	ttZ	WZ	ttX	Non-promt	Rare
nlepton = 3	781.5 ± 10.6	13233.1 ± 84.4	2195.9 ± 41.8	1982901.2 ± 1406.5	122299.4 ± 1826.5
$lepton p_T > 40, 20, 10$	299.5 ± 3.0	6863.8 ± 18.2	378.2 ± 3.0	4144.8 ± 46.1	6772.2 ± 95.5
$\|m(ll) - mZ\| < 10$	231.1 ± 2.6	5764.4 ± 16.4	189.7 ± 1.4	2978.8 ± 36.2	2349.2 ± 39.8
n jets ≥ 2	217.5 ± 2.5	1001.5 ± 8.1	127.3 ± 1.1	327.0 ± 11.5	309.4 ± 18.9
n b jets ≥ 1	161.4 ± 2.2	75.0 ± 2.0	85.0 ± 0.9	96.1 ± 5.8	27.6 ± 3.4

Table 6.9: The number of signal and background events, obtained from the MC simulations and data for the non-prompt background, after each cut applied consequently.

	n bjets = 0	n bjets = 1	n bjets ≥ 2
WZ	62 %	28 %	15 %
$t\bar{t}X$	3 %	24%	50 %
Non-prompt	15 %	38 %	25%
Rare	20%	10%	10 %

Table 6.10: The contributions of the each background in b-tagged jet multiplicities used for signal extraction.

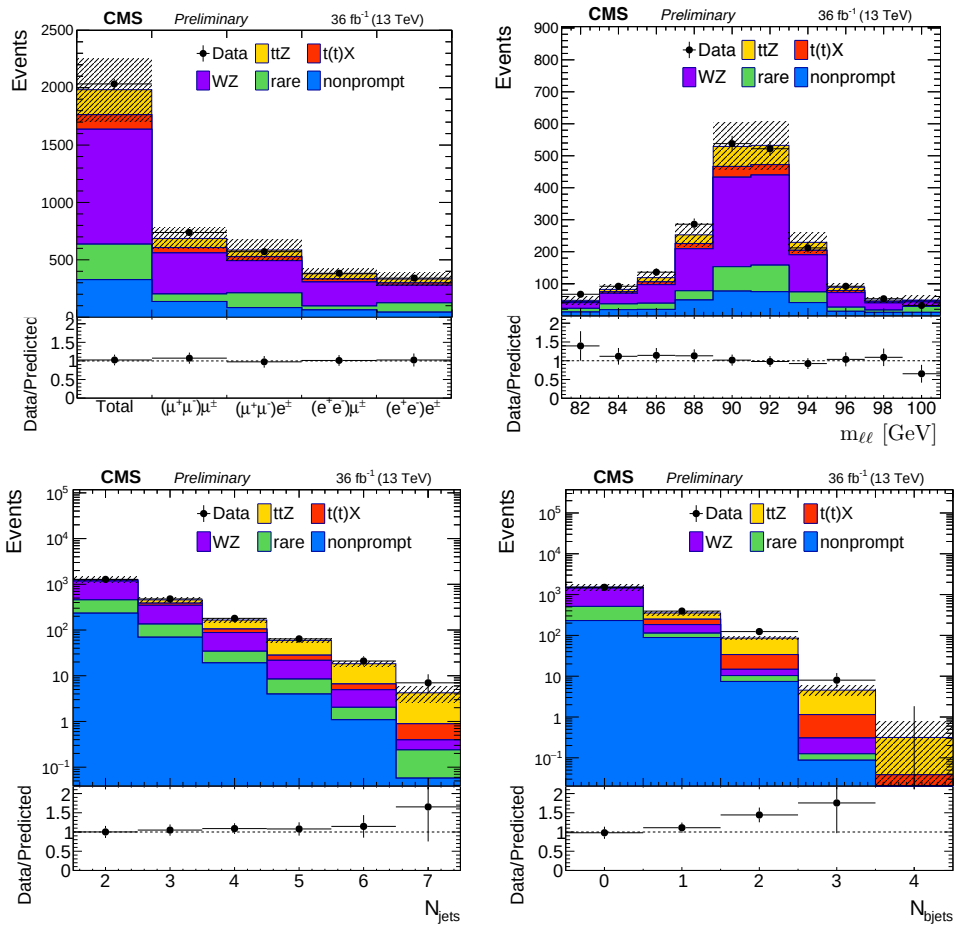

Figure 6.22: The distributions of the observed and predicted signal and background yields for events containing at least two jets. From left to right: the lepton flavor and jet multiplicity (upper), invariant mass of the SFOC lepton pair of the reconstructed Z boson, number of jets (bottom) and number of b-tagged jets. The hatched area represents the combined statistical and systematic uncertainties in the prediction.

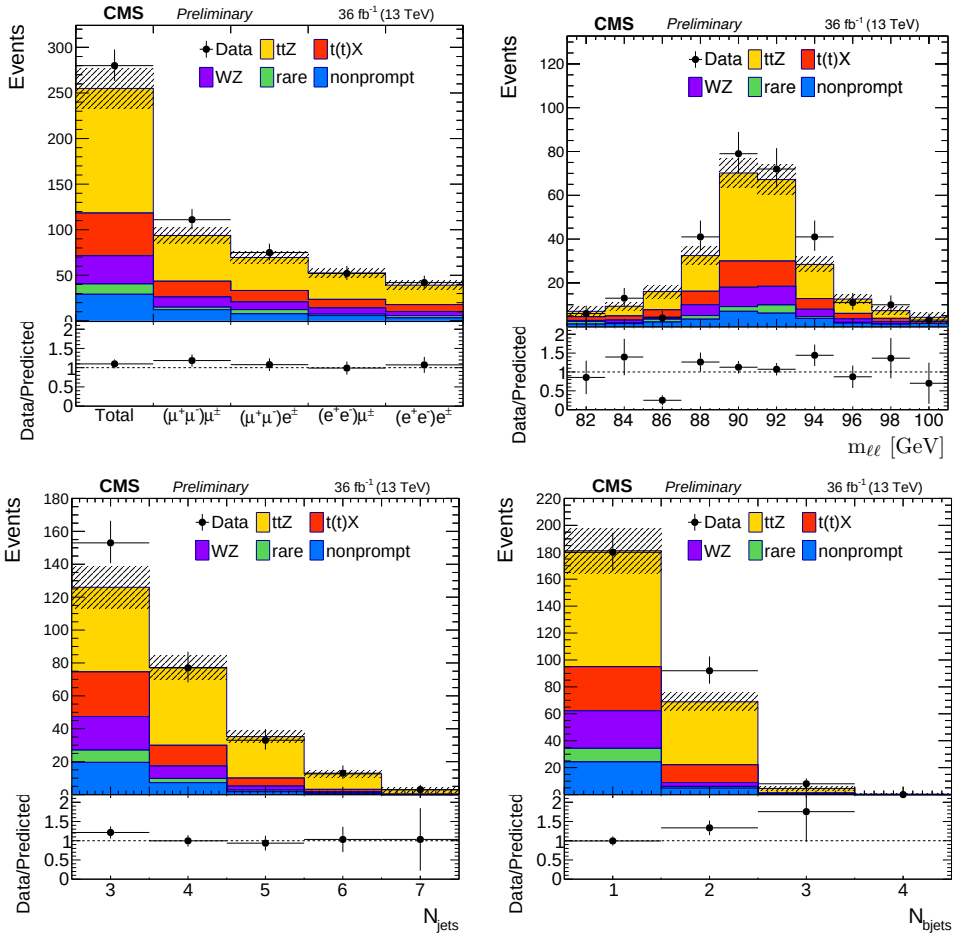

Figure 6.23: The distributions of the observed and predicted signal and background yields for events containing at least one b jet and three jets. From left to right: the lepton flavor and jet multiplicity (upper), invariant mass of the SFOC lepton pair of the reconstructed Z boson, number of jets (bottom) and number of b-tagged jets. The hatched area represents the combined statistical and systematic uncertainties in the prediction.

	2 jets		
Process	0 bjet	1 bjet	$\geq$ 2 bjet
Background	1071.1 $\pm$ 178.3	151.0 $\pm$ 23.4	10.5 $\pm$ 1.9
WZ	674.8 $\pm$ 135.1	42.6 $\pm$ 8.6	1.5 $\pm$ 0.4
ttX	23.1 $\pm$ 2.8	29.5 $\pm$ 3.6	4.7 $\pm$ 0.6
Non-prompt	166.2 $\pm$ 50.5	64.4 $\pm$ 19.9	2.7 $\pm$ 1.3
Rare	207.0 $\pm$ 104.8	14.6 $\pm$ 7.8	1.7 $\pm$ 1.1
ttZ	15.2 $\pm$ 1.6	19.6 $\pm$ 2.0	5.2 $\pm$ 0.7
Predicted	1086.3 $\pm$ 178.3	170.5 $\pm$ 23.5	15.7 $\pm$ 2.0
Data	1034	214	32

Table 6.11: The predicted and observed yields in 2 jets bin. The error represents the combined statistical and systematic uncertainties.

	3 jets		
Process	0 bjet	1 bjet	$\geq$ 2 bjet
Background	310.6 $\pm$ 51.4	60.9 $\pm$ 8.1	10.5 $\pm$ 1.4
WZ	191.8 $\pm$ 38.6	18.8 $\pm$ 3.9	1.6 $\pm$ 0.4
ttX	11.1 $\pm$ 1.4	17.9 $\pm$ 2.2	6.1 $\pm$ 0.8
Non-prompt	50.0 $\pm$ 15.8	17.2 $\pm$ 5.6	2.4 $\pm$ 1.1
Rare	57.8 $\pm$ 30.2	7.0 $\pm$ 3.7	0.5 $\pm$ 0.3
ttZ	18.8 $\pm$ 2.1	35.8 $\pm$ 3.6	15.6 $\pm$ 1.8
Predicted	329.4 $\pm$ 51.5	96.7 $\pm$ 8.8	26.2 $\pm$ 2.3
Data	326	107	46

Table 6.12: The predicted and observed yields in 3 jets bin. The error represents the combined statistical and systematic uncertainties.

	$\geq$ 4 jets		
Process	0 bjet	1 bjet	$\geq$ 2 bjet
Background	96.7 $\pm$ 15.8	29.9 $\pm$ 3.9	10.1 $\pm$ 1.5
WZ	59.8 $\pm$ 12.1	9.1 $\pm$ 2.0	1.5 $\pm$ 0.4
ttX	5.2 $\pm$ 0.7	10.7 $\pm$ 1.3	5.4 $\pm$ 0.7
Non-prompt	14.7 $\pm$ 5.0	7.1 $\pm$ 2.6	2.5 $\pm$ 1.2
Rare	17.0 $\pm$ 8.8	3.1 $\pm$ 1.7	0.8 $\pm$ 0.6
ttZ	22.1 $\pm$ 2.6	50.4 $\pm$ 5.1	35.0 $\pm$ 3.8
Predicted	118.8 $\pm$ 16.0	80.3 $\pm$ 6.4	45.1 $\pm$ 4.1
Data	147	73	54

Table 6.13: The predicted and observed yields in $\geq$ 4 jets bin. The error represents the combined statistical and systematic uncertainties.

Chapter 7

Measurement of the $t\bar{t}Z$ Cross Section

7.1 Introduction

The last part of the $t\bar{t}Z$ analysis is the cross section measurement of this process in three lepton final state. Up to now, we have established the event yields for the signal, the background and also for the data yields. Therefore we would like to compare the data to our signal and background predictions with the use of statistical methods described in this chapter. In order to do this, firstly the uncertainties associated with the expected signal and predicted background yields need to be introduced and will be followed by the statistical methods. Lastly the results will be presented.

7.2 Systematic Uncertainties

The quantification of systematic uncertainty of a measurement is not straightforward and may not have a unique definition. The uncertainties which are not statistical can generally be categorised as systematic uncertainties. While a more precise definition could be as a possible unknown variation or uncertainty in a measurement, or in a quantity derived from a set of measurements, that does not randomly vary from data point to data point. The systematics uncertainties arise from various sources such as calibration of the detectors, probability of detection of a given type of interaction or parameters of the model used for simulation. In the following sections the systematic uncertainties effecting this analysis will be discussed.

7.2.1 Experimental Uncertainties

Luminosity

A source of systematic uncertainty comes from the measurement of the integrated luminosity. This directly affects the estimation of number of events based on MC simulations which is obtained by normalising the number of events with respect to the integrated luminosity. The uncertainty on the measurement of the integrated luminosity of the 2016 data is set to 2.5% [112] which is the quadratic sum of the errors of the calibration and stability of the CMS detectors used for luminosity measurements. These are the main sources of the luminosity uncertainty, and explained in detail in [113, 112].

Pile Up

The total number of interactions in data is estimated from the measured luminosity multiplied by the average total inelastic cross section which leads to two uncertainties: the luminosity uncertainty (2.5%), and the uncertainty in the total inelastic proton proton collision cross section (4.6%). Therefore the total uncertainty assigned to PU calculation is set to 5%. The pile-up distributions are calculated by varying the expected minimum bias cross section from the mean value by $\pm$5%. The effect of this shift on the normalisation of the MC events is less than 2%.

Jet Energy

The main uncertainties of the jet energy correction, JEC, comes from the corrections explained in Sec. 5.2.4. The four vector momentum of the jet is varied up and down by using the uncertainty of the JEC as a function of the jet transverse momentum and jet pseudo rapidity. This can affect the kinematic properties of each jet in the event, which leads to a different number of jets passing or failing the selection requirements. The final effect on the event yield varies between 1 to 4% for different jet multiplicities.

The uncertainty on jet energy resolution, JER, is calculated in a similar way as JEC and has an effect of 1 to 5% for different jet multiplicities.

B-tagging

As it is explained in Sec. 6.6.4, scale factors are applied in order to compensate the b-tagging performance between data and simulation. These scale factors have associated uncertainties which are the source of b-tagging systematics uncertainty. The flavour[1] , p_T and η dependent scale factors are varied up and down with the uncertainties and the effect is propagated

[1]Here flavour refers to the mother parton of the jet, either light quarks (u,d,s) an gluons or heavy quarks (c,b)

to the event yields. The final effect on the event yield for $t\bar{t}Z$ signal and background processes vary between 1 to 5% for different b-tagged jet multiplicities.

Lepton Identification and Trigger Efficiency

Another source of systematic uncertainty comes from the lepton identification and trigger efficiency. The related corrections applied to simulation are explained in Sec. 6.6.1, 6.6.2 which have their corresponding uncertainties. In case of electron identification, the overall uncertainty of the scale factor is applied as a function of the lepton η and p_T for electrons with p_T greater than 10 GeV. While for muons, the uncertainties are calculated separately for different sources. These are 1% for the identification of muons with p_T greater than 20 GeV and 3% for muons with p_T smaller than 20 GeV. The different uncertainties used for different p_T ranges used for the identification of muons is due to the fact that the tag and probe method applied by using the Z boson invariant mass shape can not be applied for muons with low momentum $<$ 20 GeV (more correctly, very challenging due to large QCD background in the tag-and-probe data sample). Therefore for the uncovered p_T range a larger uncertainty is assigned. In addition to the muon identification a 0.5% uncertainty is applied for the muon isolation SF's. For the final muon uncertainty, they are quadratically summed. The overall effect on the event yield is about 5%.

The trigger efficiencies measured in simulation are in overall good agreement with the ones measured in data within 1%, with an exception when it is calculated as a function of the trailing lepton p_T, where the disagreement between simulation and data reaches up to 3%, Fig. 6.18. Therefore, conservatively, a 3% systematic uncertainty is assigned to trigger efficiency.

7.2.2 Theoretical Uncertainties on Signal

The sources of the theoretical uncertainties on signal are due to the parameters used for the calculation of the cross section or event yield. The factorisation, renormalisation scale uncertainties and PDF uncertainties need to be propagated to the event yields. The uncertainties of μ_R, μ_F are saved in the simulated samples as set of weights and varied up and down by a factor of two from their default values. The effects on the signal yield is found to be less than 1%. Since the PDFs are determined experimentally with some errors that come from experimental measurements, theoretical models used to extract PDFs, inclusion of any PDF into the cross section calculation introduces an additional systematic error. The PDF uncertainties are calculated by the procedure recommended by the PDF4LHC group [114]. The corresponding acceptance uncertainty amounts to less than 1%.

7.2.3 Background Uncertainties

The uncertainties assigned to each background is different depending on the theoretical uncertainty of the cross sections used for the normalisation of the simulated events and on the method used for the estimation of the background. The uncertainties assigned to the backgrounds, explained in Sec. 6.5, are as follows.

The theoretical uncertainty on the cross sections for top quark (pair) production in association with a Higgs boson or a vector boson (categorised as $t\bar{t}X$) is 11% [115]. The sources of this uncertainty comes from the choices of the scales and PDFs, as explained in the previous section. In addition to the cross section uncertainty all the experimental uncertainties, mentioned in the previous section, are applied to the $t\bar{t}X$ background.

The WZ background is studied in a data control sample, enriched by this process. In order to do this, two different NLO MC simulations were used for the WZ sample: MadGraph MG5_aMC@NLO and POWHEG. A good agreement between the data and prediction is observed for both samples for events with $N_{\text{jets}} < 4$ while the agreement is poorer for $N_{\text{jets}} \geq 4$ events for the POWHEG sample. Therefore a 10% uncertainty for the $N_{\text{jets}} < 4$ events and a 20% uncertainty for the $N_{\text{jets}} \geq 4$ are assigned in order to account this difference. In addition, the jet selection related experimental systematics (JER, JES, b-tagging) are applied on top of the cross section uncertainty to account for the discrepancy seen in high jet multiplicity events.

For the rare SM processes a 50% uncertainty is assigned. This is due to the fact that some of these processes have not yet been measured and some of the cross section values are not calculated at NLO. The additional experimental systematic uncertainties in these processes are not applied due to the fact that their contribution would be negligible with respect to the 50% uncertainty assigned to their normalisation.

Finally the statistical uncertainty due to the limited number of simulated events in MC simulation samples is considered for all the backgrounds. This ranges between 1 to 15% for $t\bar{t}X$ and WZ, and 15 to 60% for rare processes.

The uncertainties applied on background processes with non-prompt lepton(s) differ from those explained above. This is due to the fact that the backgrounds with non-prompt lepton(s) are estimated using the data. A 30% systematic uncertainty is assigned for the non-prompt background. This value comes from the discrepancy between the predicted and observed yields/distributions in simulated events, Sec. 6.5.1, that can be up to 30% in some bins. An additional statistical uncertainty due to the limited number of observed data events, used together with the fake-rate to estimate the this background, is also taken into account. Depending on the analysis category this uncertainty varies between 10 to 40%.

7.3 Statistical Analysis of $t\bar{t}Z$ Cross Section Measurement

The statistical procedures used to interpret the data and the results are governed by the LHC Higgs combination group and are explained in [116]. In this section the methods used to extract the cross section and the calculation of the signal significance in he background-only hypothesis will be explained.

7.3.1 Signal Extraction Using the Maximum Likelihood Fit

Our goal is to determine the amount of signal events to be added to the predicted background events to provide the best description of the data. The expected number of signal events are denoted by s and background events by b while the data by n_{obs} and is described as:

$$n_{obs} = \mu s + b \tag{7.1}$$

where μ is the signal strength modifier. It can have different values to represent various hypotheses of the signal: 0 for no signal or *background-only* hypotheses, 1 for the SM prediction or *signal + background* hypotheses. The probability of a given hypothesis to be represented by the data is described by the likelihood function, where the expected number of event yields are distributed according to a Poisson distribution. Thus both signal and background yields are subject to multiple uncertainties so that the signal and background yields become functions of these uncertainties (also called as nuisance parameters) denoted by θ. The corresponding likelihood function is:

$$\mathcal{L}(data|\mu,\theta) = \text{Poisson}(\text{data}|\mu \cdot s(\theta) + b(\theta)) \cdot p(\tilde{\theta}|\theta)) \tag{7.2}$$

where $p(\tilde{\theta}|\theta)$ is the systematic error probability density function (pdf) to reflect the degree of belief on what the true value of θ could be and $\tilde{\theta}$ is the default value of the nuisance parameter. In this analysis the pdf's of the nuisances are taken to be log normally distributed. Poisson(data$|\mu \cdot s + b$) stands for a product of Poisson probabilities to observe n_i events in bins i:

$$\text{Poisson}(n_i|\mu \cdot s + b) = \prod_i \frac{(\mu \cdot s_i + b_i)^{n_i}}{n_i!} e^{-\mu \cdot s_i - b_i}. \tag{7.3}$$

In order to have the best description of the data in terms of the signal strength modifier and the nuisance parameters, the Maximum Likelihood (ML) fit is used. The likelihood is maximised with respect to μ and θ. The values of μ and θ that maximise the $\mathcal{L}$ are called best fit values and denoted by $\hat{\mu}$ and $\hat{\theta}$. The most convenient way to to maximise the likelihood is to use the logarithm of the $\mathcal{L}$ where the products of the $\mathcal{L}$ are converted into summation abling an easier computation. The maximum of the likelihood is obtained by using a numerical minimisation program called MINUIT [117].

7.3.2 The Profile Likelihood as Test Statistics

Once we obtain the signal strength parameter, we would like to determine the agreement of the data with the prediction. The prediction is the outcome of a given hypothesis H which is designated as *signal + background*. H is tested against a null hypothesis H_0 which is designated as *background-only*. In order to test a hypothesised value of the signal parameter a profile likelihood ratio is defined as:

$$\hat{t}_\mu = -2\ln\lambda = -2\ln\frac{\mathcal{L}(\text{data}|\mu,\hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta})}. \tag{7.4}$$

In the numerator the nuisances are fitted to their values $\hat{\theta}_\mu$ which maximises the likelihood for a given value of μ while $\hat{\mu}$ and $\hat{\theta}$ in the denominator define the global maximum of the likelihood as explained in the previous section. According to the Wilk's theorem [118], the test statistic $\hat{t}_\mu$ asymptotically distributed as χ^2 distribution for one degree of freedom under *signal + background* experiments. In the asymptotic limit the Likelihood function becomes a Gaussian centered around the maximum likelihood estimator $\hat{\mu}$, and the profile likelihood curve has a parabolic shape:

$$-2\ln\lambda(\mu) = -2(\ln L(\mu,\hat{\theta}_\mu) - \ln L(\hat{\mu},\hat{\theta})) = \left(\frac{\mu-\hat{\mu}}{\sigma_\mu}\right)^2 \tag{7.5}$$

where σ_μ represent the Gaussian standard deviations of the parameter μ. Therefore the best fit value of the signal strength modifier can be extracted by minimising the test statistic. The 1 σ interval of the μ is obtained at the value where $-2\ln\lambda(\mu)$ is equal to unity in the profile likelihood curve [119]. This is shown for the tt̄Z analysis case in Fig. 7.2.

From the definition of λ in Eq. 7.4, the value of λ near 1 implies good agreement between the data and the hypothesised value of the signal strength parameter. Therefore higher values of t_μ correspond to increasing incompatibility between the data and μ. In order to quantify the level of dis-agreement the p-value is computed and defined as:

$$p_\mu = \int_{t_{\mu,\text{obs}}}^{\infty} f(t_\mu|\mu)dt_\mu, \tag{7.6}$$

where $t_{\mu,obs}$ is the value of the test statistic observed from the data and $f(t_\mu|\mu)$ denotes the probability density function of t_μ under the assumption of the signal strength μ [120].

In order to state the presence of a signal as a discovery, the p-value is quantified with rejecting the *background-only* hypothesis, which is $\mu = 0$. The

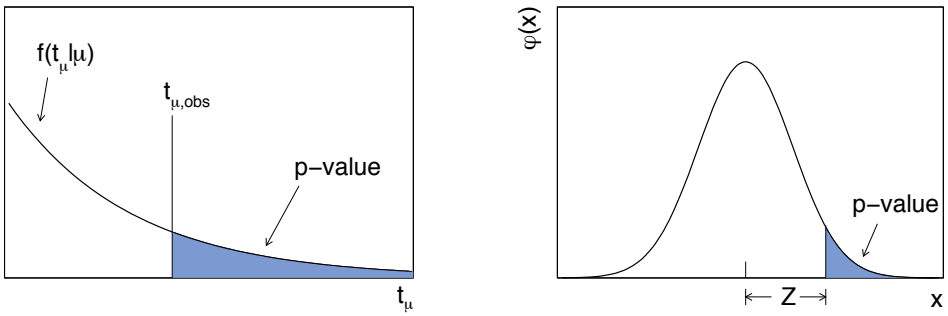

Figure 7.1: Left: Illustration of the relation between the p-value obtained from an observed value of the test statistic t_μ. Right: The standard normal distribution showing the relation between the significance Z and the p-value [120].

corresponding p-value is computed from the test statistic t_0, which is defined as:

$$t_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases}, \tag{7.7}$$

where $\lambda(0)$ is the profile likelihood ratio for $\mu = 0$ as defined in Eq. 7.4. To quantify the level of disagreement between the data and the hypothesis of $\mu = 0$ using the observed value of t_0, the p_0-value is computed in the same manner as in Eq. 7.5. This information can also be expressed in terms of the significance, Z, which is defined such that a Gaussian distributed variable is found Z standard deviations above its mean and has an upper-tail probability equal to p_0. The relation between the p-value and the observed t_μ and also the significance Z are shown in Fig.7.1 [120]. The 3 standard deviation (σ) corresponding to a p_0 value of 0.00135 is accepted as an evidence of observing a signal while 5σ is qualified as a discovery with a corresponding p-value of 2.87×10^{-7}. In 5σ case, the probability of the background processes to produce events that could correspond to the signal is 0.0000287% [32].

The p-value can also be used for excluding the signal hypothesis if the compatibility with the *signal + background* hypothesis is small. This is called as exclusion limit-setting and mostly used for BSM searches to exclude a new signal. The procedure is based on a statistic called $\mathrm{CL_s}$ which is defined as:

$$\mathrm{CL_s} = \frac{p_\mu}{1 - p_\mathrm{b}}, \tag{7.8}$$

where p_b is the p-value of the *background-only* hypothesis [32]. A new signal or a point in a model's parameter space is regarded excluded if one finds $\mathrm{CL_s} \leq \alpha$, then it is excluded with $(1 - \alpha)$ $\mathrm{CL_s}$. A common choice of α is 5% which corresponds to an exclusion limit of 95% confidence level. This method is called as the modified frequentist method [121].

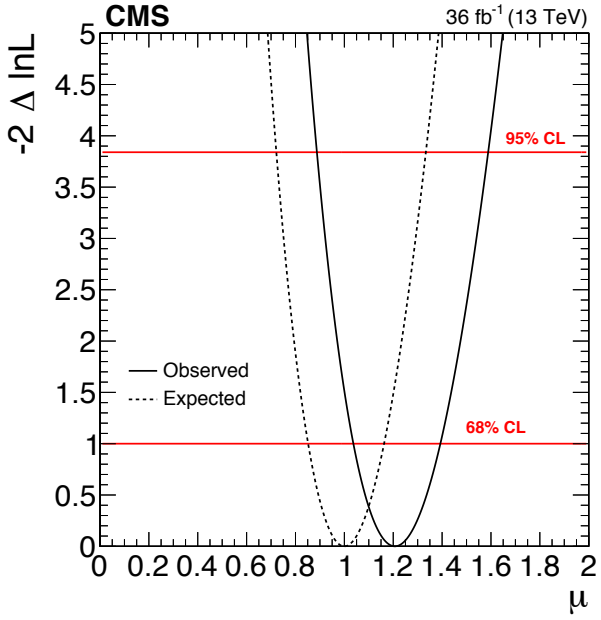

Figure 7.2: The expected and observed signal strength modifier with their 68 and 95% CL uncertainties.

7.4 Results

7.4.1 3-lepton Final State

As described in Section 6.4, the analysis is made in several jet and b-tagged jet multiplicities as given below in the 3 lepton final state:

- 2 jets and 0, 1, $\geq$2 b-tagged jets;
- 3 jets and 0, 1, $\geq$2 b-tagged jets;
- $\geq$ 4 jets and 0, 1, $\geq$2 b-tagged jets.

The event selections described in Sec. 6.4 are applied to the 9 analysis bins summarized above. For the signal, the simulations and for the background, either simulations or data are used. The cuts determined to select the tt̄Z events are determined and then applied to the signal and background processes without looking to the data distributions, in other words the data in the signal region was blinded until all systematic effects were studied and concluded. Before unblinding, the data is only used in control regions. The final step of the analysis is to unblind the data, meaning applying the event selections to the data and extracting the observed distributions and finally the results. The data used in this analysis is collected by CMS during 2016 and corresponds to an integrated luminosity of 36 fb^{-1}.

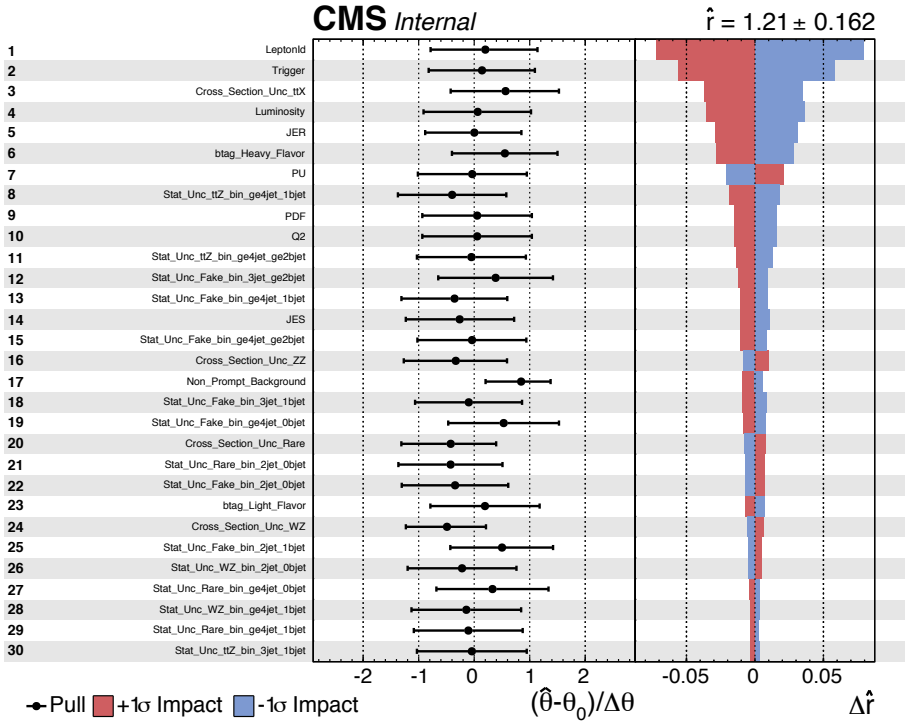

Figure 7.3: Summary of the sources of uncertainties sorted by decreasing impact on the result, right column. On the left column of the plot the pull distribution of the nuisances are shown.

The target of the analysis is to measure the cross section of the $t\bar{t}Z$ process with its statistical significance. In order to extract the the signal strength modifier μ with its associated 68% CL interval, a likelihood fit and its asymptotic approximation are performed in the defined 9 analysis bins. The value of the best fit of the signal strength modifier that describes the data is found to be:

$$\mu = 1.21^{+0.11}_{-0.11}(\text{stat.})^{+0.14}_{-0.12}(\text{sys.})^{+0.13}_{-0.14}(\text{theo.}) \tag{7.9}$$

with an observed (expected) significance of 9.77 (9.04). The expected and observed signal strength modifier as a function of the test statistics used to obtain its value and corresponding uncertainties are shown in Fig. 7.2. The measured signal strength modifier is compatible with its expected SM value. In order to extract the cross section of the $t\bar{t}Z$ production, the cross section value used for the normalisation of the signal sample is multiplied with the signal strength modifier. As given in Chapter 6, Tab. 6.2 the NLO cross section of $t\bar{t}Z/\gamma^* \rightarrow t\bar{t}(\ell\ell, \nu\nu)$ is 0.2728 pb. This needs to be corrected for the branching ratio of $Z\gamma^*$ decaying to charged leptons $(\ell\ell)$ and Z to neutral leptons $(\nu\nu)$, which is $3 \times 0.0336 \times 1.24 + 0.20$ [32]. The first value accounts for the branching ratio of Z/γ^* to charged leptons[2] and the second to neu-

[2]Here the factor of 1.24 accounts for the effect of γ^*. This value is obtained by the

trinos in percentages. Therefore the cross section of $t\bar{t}Z$ in all decay modes is found to be 0.839 pb. This value is in agreement with the theoretical cross section computed at NLO using MadGraph MG5_aMC@NLO which is 0.839±0.101 pb [115]. Finally the measured cross section of $t\bar{t}Z$ extracted from the 3-lepton final state is found to be:

$$\sigma(pp \to t\bar{t}Z) = 1,01^{+0.09}_{-0.09}(\text{stat.})^{+0.10}_{-0.11}(\text{sys.})\text{pb}. \tag{7.10}$$

The systematic uncertainties explained in Sec. 7.2, that are set as free parameters in the fit, do not have the same effect on the results. Therefore the impact of each systematic uncertainty on μ is defined as the shift of $\Delta\mu$ when θ is fixed and brought to its $\pm 1\sigma$ post-fit values while other systematics are treated as usual :

$$\Delta\mu(\pm) = \hat{\mu}(\hat{\theta} \pm \Delta_\theta) - \hat{\mu}(\hat{\theta}). \tag{7.11}$$

The impact of each systematic source in ascending order are shown in Fig. 7.3. The uncertainties related with lepton identification and trigger have the largest effect on the result. In addition to the impact of the uncertainties, the best fit value of the nuisances after the fit (post-fit) may be different from their initial (pre-fit) values given as a free parameter in the fit. Some of the nuisances may be pulled up (down) having larger (smaller) post-fit values while some may not get constrained in the fit. This difference can be described as $(\hat{\theta} - \theta_0)/\Delta\theta$ where $(\hat{\theta}, \theta_0)$ are the post-fit and the pre-fit values of θ respectively and $\Delta\theta$ is its uncertainty. The pull distribution of the nuisances are shown in Fig. 7.3, where it can be seen that the uncertainties due to the cross section uncertainty of $t\bar{t}X$, btagging, are pulled up and some statistical uncertainties are pulled up and down by less than 1 sigma. While the uncertainty on the non-prompt background is pulled up and also has a lower post-fit value, which can be seen from the width of the pull. This is a result of the fact that some analysis categories with substantial non-prompt lepton contribution has a very good data-prediction agreement such that the data suggests a smaller uncertainty overall.

Finally the post-fit distribution of the analysis bins is shown in Fig. 7.4 where all the yields and systematic uncertainties are the post-fit one. The corresponding number of events in each jet multiplicity bins are shown in Tables 7.1-7.3. In general a good agreement is seen with the predicted and observed yields except for some excess of events in $N_{\text{jets}} = 2, 3$ and $N_{\text{b jets}} = 2$ categories. Extensive studies were performed to ensure this is not due to any missing or underestimated backgrounds and no related indication is found. Therefore, this excess is attributed as a statistical fluctuation in data.

comparing the LO cross sections of $t\bar{t}Z$ and $t\bar{t}Z/\gamma^*$.

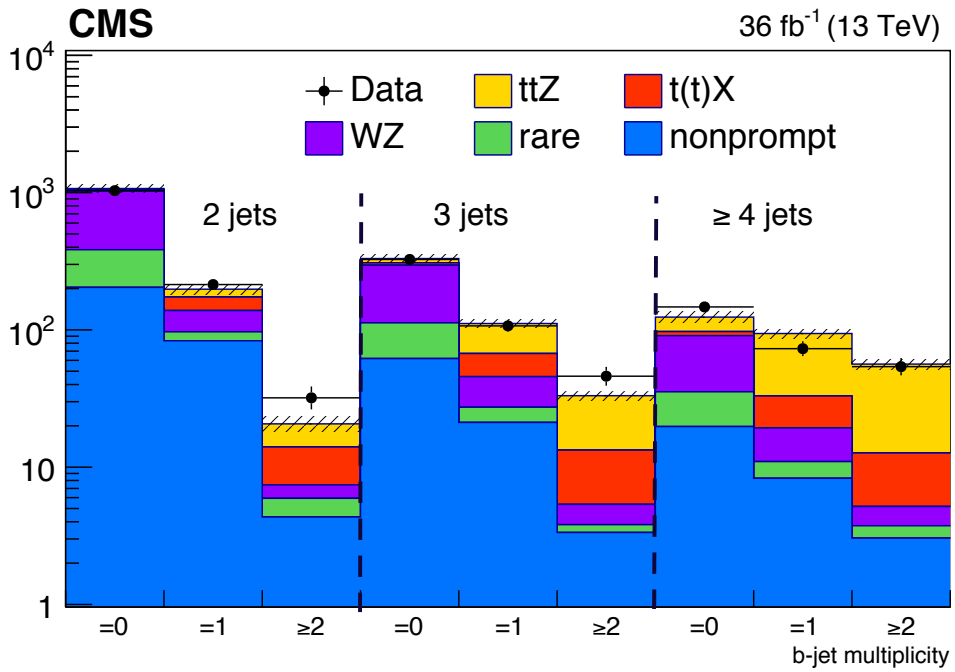

Figure 7.4: Predicted signal and background yields, as obtained from the fit, compared to observed data in $N_{\text{jets}} = 2$, 3 and $\geq$4 jets categories with $N_{\text{b jets}} = 0$,1and ≥ 2 b-jet multiplicities. The hatched band shows the total uncertainty associated to signal and background predictions..

	2 jets		
Process	0 bjet	1 bjet	$\geq$ 2 bjet
Background	1052.9 $\pm$ 68.4	173.6 $\pm$ 15.7	14.0 $\pm$ 2.5
WZ	643.2 $\pm$ 44.6	41.5 $\pm$ 3.9	1.5 $\pm$ 0.3
ttX	25.9 $\pm$ 3.0	35.2 $\pm$ 4.0	6.6 $\pm$ 0.7
Non-prompt	204.0 $\pm$ 32.7	83.4 $\pm$ 13.7	4.3 $\pm$ 2.0
Rare	179.7 $\pm$ 40.3	13.5 $\pm$ 5.2	1.6 $\pm$ 1.2
ttZ	18.4 $\pm$ 2.1	24.4 $\pm$ 2.8	6.6 $\pm$ 0.8
Predicted	1071.4 $\pm$ 68.4	198.0 $\pm$ 15.7	20.7 $\pm$ 2.5
Data	1034	214	32

Table 7.1: Predicted signal and background yields, as obtained from the fit, compared to observed data in 2 jets bin.

	3 jets		
Process	0 bjet	1 bjet	$\geq$ 2 bjet
Background	308.3 $\pm$ 23.4	67.3 $\pm$ 5.6	13.3 $\pm$ 1.6
WZ	183.1 $\pm$ 13.7	18.3 $\pm$ 1.8	1.5 $\pm$ 0.3
ttX	12.7 $\pm$ 1.5	21.7 $\pm$ 2.3	8.0 $\pm$ 0.9
Non-prompt	61.8 $\pm$ 12.0	21.2 $\pm$ 4.3	3.3 $\pm$ 1.3
Rare	50.7 $\pm$ 14.5	6.2 $\pm$ 2.1	0.5 $\pm$ 0.1
ttZ	22.7 $\pm$ 2.8	44.0 $\pm$ 4.5	19.8 $\pm$ 2.2
Predicted	331.0 $\pm$ 23.4	111.3 $\pm$ 5.6	33.2 $\pm$ 1.6
Data	326	107	46

Table 7.2: Predicted signal and background yields, as obtained from the fit, compared to observed data in 3 jets bin.

	$\geq$ 4 jets		
Process	0 bjet	1 bjet	$\geq$ 2 bjet
Background	97.3 $\pm$ 10.6	33.0 $\pm$ 3.1	12.7 $\pm$ 1.6
WZ	55.4 $\pm$ 7.8	8.3 $\pm$ 1.4	1.4 $\pm$ 0.3
ttX	6.5 $\pm$ 0.8	13.7 $\pm$ 1.4	7.5 $\pm$ 0.8
Non-prompt	19.8 $\pm$ 4.8	8.3 $\pm$ 2.1	3.0 $\pm$ 1.2
Rare	15.6 $\pm$ 5.3	2.7 $\pm$ 1.0	0.7 $\pm$ 0.4
ttZ	26.7 $\pm$ 3.4	61.1 $\pm$ 6.3	43.7 $\pm$ 4.8
Predicted	124.0 $\pm$ 10.6	94.1 $\pm$ 3.1	56.4 $\pm$ 1.6
Data	147	73	54

Table 7.3: Predicted signal and background yields, as obtained from the fit, compared to observed data in $\geq$ 4 jets bin.

7.4.2 The Combination of 3 and 4 Lepton Final States

In this thesis, the $t\bar{t}Z$ production is studied in the 3-lepton final state. This result is combined with the $t\bar{t}Z$ production in the 4-lepton final state where the top-anti-top pair and the Z boson decay into leptons. The 4-lepton analysis is performed with the selection of exactly 4 leptons where at least one SFOC lepton pair with an invariant mass of $|m_{\ell\ell} - m_Z|$ GeV and at least 2 jets are present in the event. For a detailed description of the event selections of the 4-lepton analysis see [8]. Events are categorised according to their b-tagged jet multiplicities: 0 and 1 b-tagged; where these two bins are used for the fit. In order to combine the results of 3- and 4-lepton analyses, the statistical procedure described in Sec. 7.3 is used. The common systematic uncertainties for both of the analyses are taken correlated. These are the

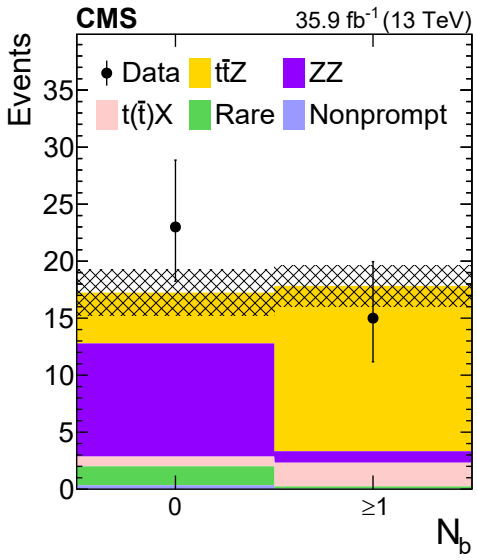

Figure 7.5: Predicted signal and background yields, as obtained from the fit, compared to observed data in $N_{bjets} = 0, \geq 1$ categories in the 4-lepton analysis. The hatched band shows the total uncertainty associated with the signal and background predictions as obtained from the fit, [8].

Process	0 bjet	$\geq$ 1 bjet
Total background	12.8±2.0	3.3±0.3
tt̄Z	4.5±0.6	14.5±1.8
Total	17.2±2.0	17.8±1.8
Observed	23	15

Table 7.4: Predicted signal and background yields, as obtained from the fit, compared to observed data in the four-lepton final state. The total uncertainty obtained from the fit is also shown, [8].

experimental and theoretical uncertainties explained in Sec. 7.2.1 and 7.2.2. The uncertainties which are not taken correlated are the statistical uncertainties of the MC samples and the estimation of non-prompt background, since in 4-lepton analysis the estimation of the non-prompt background is done with MC simulations not with the data-driven method explained in Section 6.5.1.

The predicted SM background and signal yields, and the observed data in 4-lepton final state as obtained from the fit are shown in Fig. 7.5 and in Table 7.4.2. The expected significance for tt̄Z in the 4-lepton final state is 4.7 while the expected one is 4.5 standard deviations.

The measured signal strength parameter of the combined tt̄Z in 3- and 4-lepton final states is found to be $1.17^{+0.11}_{-0.11}$stat.$^{+0.14}_{-0.12}$syst.$^{+0.11}_{-0.12}$theo.This is mul-

tiplied with the cross section used for the normalisation of the MC sample, as done for 3-lepton final state, to obtain the measured cross section for $t\bar{t}Z$:

$$\sigma(pp \to t\bar{t}Z) = 0.99^{+0.09}_{-0.08}(\text{stat.})^{+0.12}_{-0.10}(\text{sys.})\text{pb}, \tag{7.12}$$

The measured cross section precision is 14% and dominated by the systematic uncertainties. Therefore the precision of the measurement will increase marginally with the addition of more data sets while the gain on the precision will be from the improvement of the systematic uncertainties.

Part II

WZ scattering in view of the CMS upgrade for HL-LHC

Chapter 8

WZ scattering in view of the CMS upgrade for HL-LHC

8.1 High Luminosity LHC

The LHC proton proton collision program has been running successfully since 2010 at centre of mass energies of 7, 8 and 13 TeV. LHC is pushing the limits of technologies to reach the highest energies and luminosities in order to enable physicists to verify the SM and go beyond it. The discovery of the only missing particle of the SM, the mass giving Higgs boson, was announced in 2012 by ATLAS and CMS collaborations. It is the first fundamental discovery of the LHC. In addition to its discovery, the decay of the new boson to the W and the Z bosons, photons and to tau-antitau, bottom-antibottom pairs and also its properties were established. Although the Higgs boson with a mass of 125 GeV behaves like the SM Higgs, measurement of some its properties are still not very precise. Additionally, many searches some of which have been mentioned in Chapter 2, have been going on at LHC with the available data. Whereas, there has not yet been any evidence of BSM.

Up to now, LHC has been running with a luminosity of at most $1\times10^{34}\text{cm}^{-2}$ s^{-1} and has delivered around an integrated luminosity of 100 fb^{-1}. In 2019 the LHC will stop taking data for maintenance and upgrades, called Long Shutdown 2 (LS2). The detailed description of the updates for LS2, Phase-I upgrade, is given in [122]. The Phase-I upgrade will facilitate LHC to deliver an integrated luminosity of 300 fb^{-1}. After reaching this value, the statistical gain in running the accelerator without a significant increase in luminosity of its design luminosity will be marginal. Therefore, to explore LHCs full capacity, the LHC machine will be upgraded to run at higher luminosities. This upgrade is called Phase-II upgrade and the project is called High Luminosity LHC (HL-LHC). The current plan of LHC and HL-

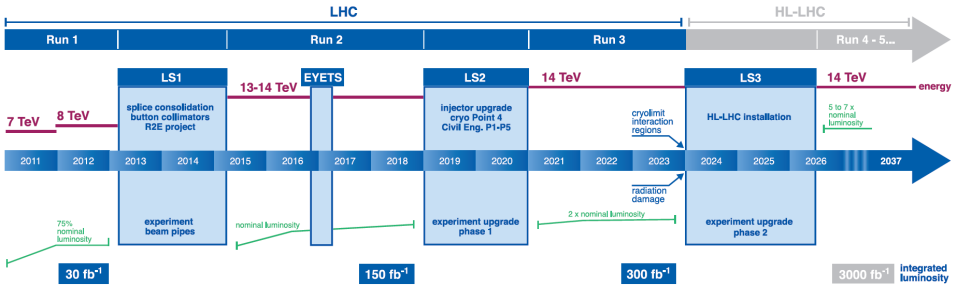

Figure 8.1: LHC plan from 2012 till now and beyond showing the energy of collisions (upper red line) and luminosity (lower green lines). Taken from [123]

LHC for a series of long periods of data-taking, referred to as Run I, Run II, Run III etc., interleaved with long shutdowns, designated LS1, LS2 and LS3 are shown in Fig. 8.1.

As its name suggests, the main objective of the HL-LHC is to run at a higher luminosity, which will reach 5-7$\times 10^{34}$ cm$^{-2}s^{-1}$, delivering around 250 fb^{-1} per year for 10 years of operation. Due to the high luminosity environment, one of the main challenges of the HL-LHC will be the high pile-up and radiation damage. In order to achieve higher luminosity values, one of the main upgrades will be in the beam injection system at Proton Synchrotron. This will happen in LS2, so that after this upgrade LHC will run at its twice nominal luminosity. By 2024, the quadruple magnets that focus the beams at the experiment collision points are expected to be close to their lifetimes due to radiation exposure. These will be replaced in LS3. A detailed description of the HL-LHC upgrades is given in [123].

The physics analyses will be inherited by the results from the data collected before LS3, which will be around 300 fb^{-1}. The study and precise measurements of the Higgs boson properties will continue to be one of the main physics goals of the HL-LHC. In addition, the role of the Higgs boson in the electroweak symmetry breaking will be tested in studies of the vector boson scattering processes. These measurements could also be sensitive to new physics through anomalous couplings. Additionally many BSM searches and exotica searches will still be exciting topics to be studied in HL-LHC.

8.2 CMS Detector Upgrade for the HL-LHC

The main goal of the Phase-II upgrade of the CMS detector is to maintain its excellent performance in the high density environment. The radiation caused by the high intensity particle collisions in the material of the detectors and on the electronics will cause significant damage and could result in a progressive degradation of the detector performance. The expected annual dose of the HL-LHC will be similar to the total dose delivered from the

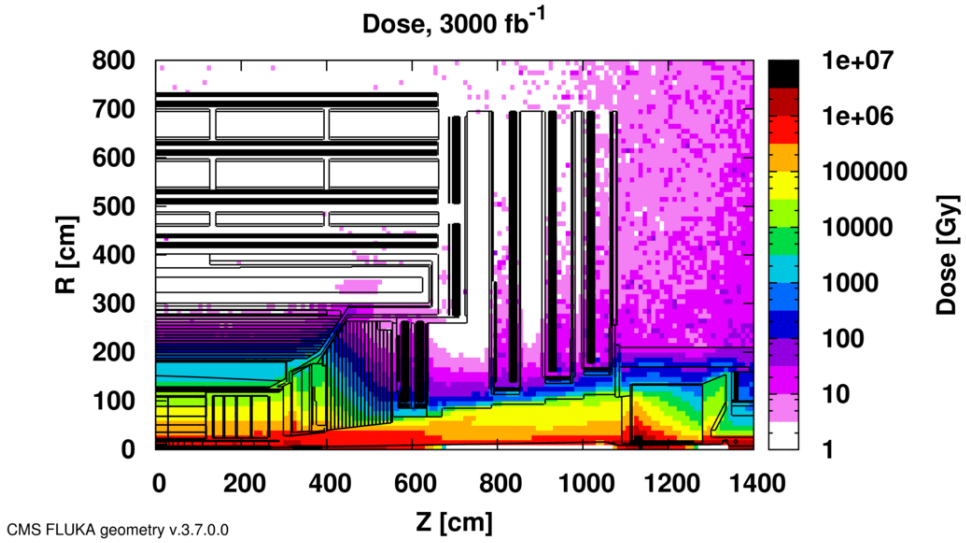

Figure 8.2: Absorbed dose in the CMS cavern after an integrated luminosity of 3 ab^{-1}. R is the transverse distance from the beamline and Z is the distance along the beamline from the Interaction Point at Z$=0$ [4].

beginning of LHC to the start of LS3. This is estimated by extrapolating the condition of the detectors after Run-I and using test beams and simulations. It is important to predict the expected dose of radiation and the detector response to radiation in order to design a detector that can maintain its performance as the integrated luminosity approaches 3 ab^{-1}. An example of the predicted radiation in CMS detector after absorption of 3 ab^{-1} integrated luminosity is shown in Fig. 8.2. As shown in the figure, the effect of the radiation dose is not the same for all components of the detector, for instance it is expected to be higher where the particle interactions/collisions are expected to be the highest due to pile-up. The components that will be affected the most by radiation will be the the tracker, the endcap calorimeters and the forward detectors, and therefore, these aged detectors need replacement for Phase-II.

In Chapter 3, as it was mentioned, the LHC proton beams are not continuous but are made of proton bunches. In order to have higher luminosities, the number of protons on each bunch needs to be increased which causes more pile-up interactions. The expected number of PU interactions in HL-LHC is around 140 which was on average between 5 to 50 in the previous runs. Therefore this significant increase in PU brings many complications in the detectors and also in the trigger and DAQ systems. A relatively high PU event was produced, in a special data taking run in 2012, with 78 reconstructed vertices. This event is shown in Fig. 8.3.

The upgrades needed for the high luminosity runs in CMS detectors and

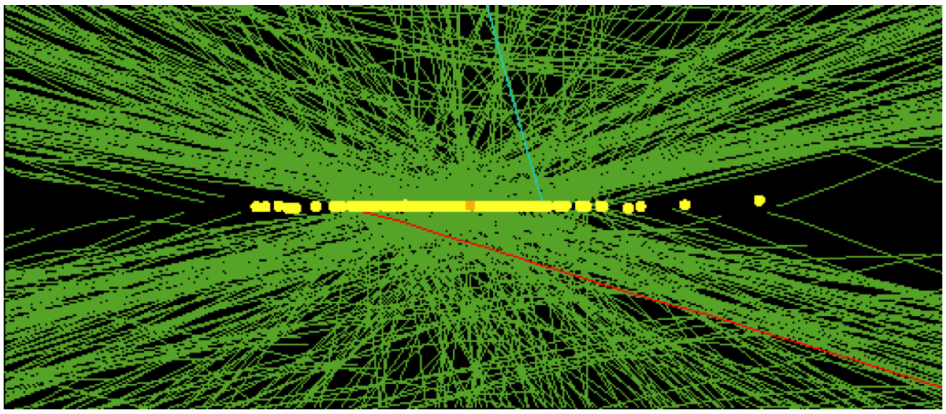

Figure 8.3: High pileup event with 78 reconstructed vertices taken in 2012. [4]

systems are briefly listed above.

- **Tracker:** The tracker will be exposed to high radiation until LS3 and need to be completely replaced for Phase-II. In order to maintain sufficient tracking performance in high PU environment the granularity of the pixel and outer strip tracker will be increased by a factor of 4 to 6. Additional design improvements will provide a better p_T resolution and will be capable of providing track information to the L1 trigger for tracks with $p_T > 2$ GeV. This is essentially needed to reject background events at the earliest stage of event selection, at L1. The inner tracking system, pixel tracker, will be implemented by smaller size pixels and thinner sensors that will result in improved impact parameter resolution and better two-track separation. This will additionally improve b-tagging and hadronic τ decay reconstruction efficiencies. In order to match with the calorimeter coverage and to improve jet reconstruction efficiency, the tracker coverage will be extended to $|\eta| \sim 4$ with addition of 10 pixel disks.

- **Calorimeters:** The barrel electromagnetic and hadronic calorimeters will maintain their performance during HL-LHC and will not be replaced. Though, the upgrade of the electronics is needed in order to satisfy the new trigger requirements. The endcap calorimeters, electromagnetic and hadronic, will suffer from high radiation dose due to PU events and need to be completely replaced for Phase-II. These will be replaced by the High Granularity Calorimeter (HGC). The HGC has electromagnetic and hadronic components that has excellent transverse and longitudinal segmentation providing a detailed three dimensional images of showers.

- **Muon system** As the barrel calorimeters, the muon system compo-

nents are expected to tolerate the increased radiation values during Phase-II without crucial degradation except the readout electronics of the DTs that will be replaced. This will improve the current trigger rate limitation of the system. Additionally for improving the trigger limits, the CSC chambers readout electronics will need to be upgraded. The major upgrade of the muon system will be the addition of four stations in the end-caps to maintain good trigger efficiency in high density conditions. The first two stations will be Gas Electron Multiplier (GEM) chambers, in this region magnetic field is still high enough that these high granularity detectors will improve position resolution. The last two stations will use RPCs with lower granularity but with good timing resolution. In addition, the extension of the muon coverage up to $|\eta| \sim 3$ is possible in the space becomes free behind the new endcap calorimeter. The insertion of a GEM station in this region is being proposed.

- **Trigger system and DAQ:** Due to very high collision rates in Phase-II, it is mandatory to implement the trigger system to have a similar performance as in Phase-I. The major upgrade in the trigger hardware system will be the addition of track information at L1. In order to allow this the trigger latency, the time needed for trigger decision, will be increased to 12.5 μs from 3.2 μs. As mentioned above, this will require upgrades in the readout systems of some detectors. Additionally, the DAQ and software systems will need to be upgraded to implement the increase of bandwidth and computing power required to accommodate the larger event size and trigger rate.

A more detailed description of the Phase-II upgrades is given in [4, 124].

8.2.1 CMS Detector Configurations

In order to study the impact of the upgrades on the physics analyses measurement capabilities at the HL-LHC, three different configurations of the detector and data taking conditions have been considered:

- **Phase-I:** the Phase-I detector that has the upgrades from LS2. It operates at 50 PU without radiation damage in order to establish as a benchmark for the expected performance of the Phase−II upgrades;
- **Phase-I aged:** the Phase-I detector, as above. But it operates at 140 PU with modelling of the effects of radiation ageing after integrated luminosities of 1000 fb^{-1}, and in a few cases up to 3000 fb^{-1}. This is also needed to determine when updates become essential;
- **Phase-II:** the Phase-II detector upgraded with the specifications explained in this chapter. It operates at 140 PU in order to evaluate the performance reach for the new concepts. It is modelled that the

performance of the new sub-detectors will not degrade with radiation while an intermediate ageing of 1000 fb^{-1} is included for the barrel calorimeters since they will not be replaced.

8.3 WZ Scattering at HL-LHC

In Section 2.4, the VBS phenomenology and its role in EWSB was introduced. Here, a dedicated analysis of the WZ scattering in view of the HL-LHC is reported. In order to asses of the sensitivity of the upgraded CMS at HL-LHC to WZ scattering the following analyses benchmarks are considered:

- **The EWK component cross-section:** With the collected amount of data at HL-LHC, the important parameter will be the precision of the measurement. The expected precision of the electroweak WZ scattering cross section is estimated.
- **The longitudinal part of the EWK component cross-section:** The expected significance of the longitudinal component of the WZ scattering cross section is studied.
- **The partial unitarization scenario:** Some BSM models predict more than one mass giving boson that breaks the spontaneous electro-weak symmetry. This presents an experimental phenomenology where the already discovered Higgs boson only partially fulfils the regularisation of the vector boson scattering. Therefore VBS measurement is studied to test the regularisation in SM and in an extreme case without a Higgs boson.

8.3.1 WZ Scattering

The scattering of W and Z bosons occurs at proton proton collisions when each of the quarks from colliding protons radiate a vector boson. These two bosons interact with each other while the two initial quarks, carrying the significant fraction of the protons energy, are deflected from their initial trajectory. The decay products of W and Z bosons and two jets originated by the incoming quarks are the final states detected in the experiment. The jets are expected to be high energetic and with large pseudo-rapidity difference with respect to each other. These jets are referred as "tag jets", while leptons from the decay products of W and Z bosons are expected to be central. A schematic drawing of the vector boson scattering is shown in Fig. 8.4.

For this study, only fully-leptonic decays of the vector bosons are considered, in order to avoid the ambiguities due to the separation of the tag jets from the decay products of the vector bosons. Only decays into electrons and muons are considered, where the contribution of τ leptons are considered by their decay to electrons and muons. The final state of WZ scattering

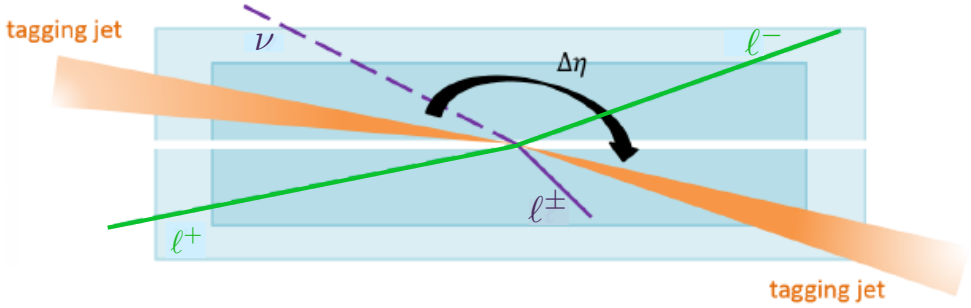

Figure 8.4: A typical signature of WZ scattering: high p_T central decay products of W and Z bosons ($\ell^\pm\nu$ and $\ell^+\ell^-$), and two hard forward well separated tagged jets, where $\Delta\eta$ shows the pseudorapidity difference of the tagged jets . Taken from [125].

is characterised by two tagged jets and three leptons where two of which have same flavour opposite charge coming from the Z boson and the third lepton coming from the W boson. The decay chain of the WZ scattering is as follows:

$$qq \rightarrow q\prime q\prime W(\ell^\pm\nu)Z(\ell^\pm\ell^\mp) \tag{8.1}$$

The signal events are produced via diagrams involving electro-weak vertices only. This is obtained with a calculation at leading order in α^6_{EWK}, which is called as the EWK component, shown in top plots in Fig. 2.9. A similar topology can result from diagrams at the order $\alpha^4_{EWK} \cdot \alpha^2_S$ corresponding to the QCD production of WZ bosons, shown in bottom plot in Fig. 2.9. These events correspond to the irreducible background and contribute as the main background. The production of ZZ bosons, decaying to four leptons with additional jets where one of the four leptons is not reconstructed, as well result in a similar final state. But it has a minor contirubtion to background with respect to the QCD production of WZ bosons. The Drell-Yan production in association with jets may contaminate the signal region when one jet is mis-identified as a lepton, but this contribution is studied and shown to be negligible with respect to the QCD WZ and ZZ production.

8.3.2 Event Generation and Reconstruction

The signal and background samples are produced with LO event generators. The MC program PHANTOM 1.2.6 [126] is used to generate events with two particles in the initial state and six particles in the final state, performing the exact calculation at leading order in $\alpha^6_{EWK} + \alpha^4_{EWK} \cdot \alpha^2_S$ for the full set of signal and backgrounds and also for the event generation of the samples where the Higgs boson only partially contributes to the unitarization.

For generating the samples for the measurement of the longitudinal part of the EWK cross-section, the polarization information of the outgoing vector bosons is needed. Therefore for this case the MADGRAPH v5.2.2.3 [83] MC program with the Decay package that provides additional information of the outgoing vector bosons, is used. All the MC programs are further interfaced with PYTHIA versions 6.4.2.6 [81] or 8.2.01 [82] for the parton shower and hadronization steps.

The different detector configurations mentioned in Sec. 8.2.1 have been implemented and processed with a detailed simulation of the interactions of particles with the detector based on GEANT4. This is followed by the event reconstruction software of the CMS experiment, explained in Chapter 5. Considering the computation time of the full simulation of the detector with GEANT4, the full simulation of the detector has been processed for a limited set of samples while all the analyses samples have been processed with a parametric implementation of the detector response based on the DELPHES [127] package. The DELPHES framework takes output of an event generator and performs a fast and realistic simulation of a general purpose collider detector. The parametrisation of efficiencies and resolutions have been tuned and implemented by comparing the DELPHES samples results with the corresponding samples produced with the full GEANT4 simulation. A more detailed study of the detector performance studies based on GEANT4 and DELPHES can be found in [4].

The online selection of the events are assumed to be collected by double-lepton triggers based on the trigger studies performed for the HL-LHC. These triggers require thresholds on the transverse momentum of the two leading leptons with $p_{\mathrm{T}} \sim 20, \sim 10$ GeV. While in the analysis all the leptons are selected with p_{T} to be greater than 20 GeV, that ensures to have a trigger efficiency of $\sim 100\%$.

The reconstruction of events, using the PF algorithm, based on the full GEANT4 simulation of the CMS is described in Chapter 5. For the Phase II detectors there are some modifications needed in the PF algorithm due to the upgrades in the detector subsystems. These are shortly the following.

- The ECAL will be replaced by the HGCal in the endcap region that rather than the clusters built in ECAL, clusters built on each layer of the HGCal are linked to form a single electromagnetic cluster. Clusters are then built as in the usual PF algorithm.
- Jet reconstruction is extended to $|\eta| < 4.7$ and lepton reconstruction to $|\eta| < 4.0$.

For the sake of this study, lepton isolation is calculated in a cone around the lepton (ΔR) of 0.3 both for electrons and muons. The residual pile-up contributions to neutral particles are subtracted with effective area corrections, as

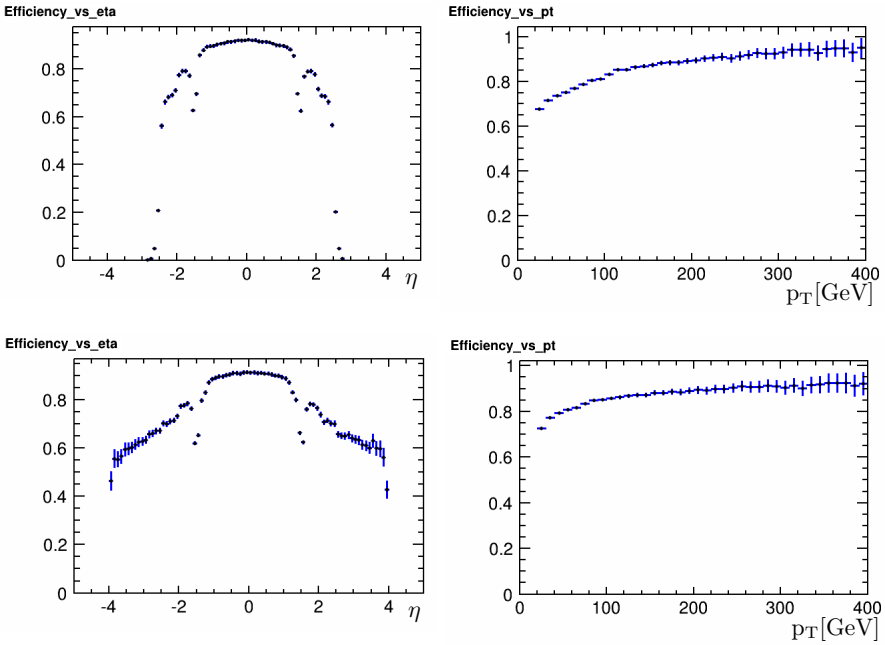

Figure 8.5: Electron reconstruction efficiency as a function of η or p_T. The Phase-I Not Aged scenario is considered on the top, the Phase-II on the bottom.

given in Formula 6.1. The isolation variable is requested to be less than 0.25 and 0.65 for Phase-I 50 PU and Phase-II 140 PU respectively. The particles surviving these isolation criteria are referred as "tight leptons" in the following. In order to veto additional leptons besides the ones expected from the decay of WZ final state, "loose leptons" are defined by loosening the isolation value to 0.35 and 0.75 for Phase-I and Phase-II respectively. The reconstruction efficiencies, identification and isolation, for tight electrons (muons) is $\sim$ 75% ($\sim$ 80%) for the barrel and drop to $\sim$ 55% ($\sim$ 65%) in the end-cap. In case of the Phase-II detector the efficiencies stay the same as in end-cap between $2.5 < |\eta| < 4.0$.

The efficiency and purity of the lepton reconstruction are determined from the full simulation and parametrically inserted into the Delphes simulation. On the contrary, the isolation selection is applied during the analysis for both the full simulation and Delphes samples. Figure 8.5, shows at the top the electron reconstruction efficiency for the configuration Phase-I-Not Aged, while at the bottom the same distributions for the configuration Phase-II , including the lepton ID. The corresponding trends for muons are reported in Figure 8.6.

Table 8.1 shows the list of generated samples, together with the generator used, the total number of produced events and the corresponding cross-

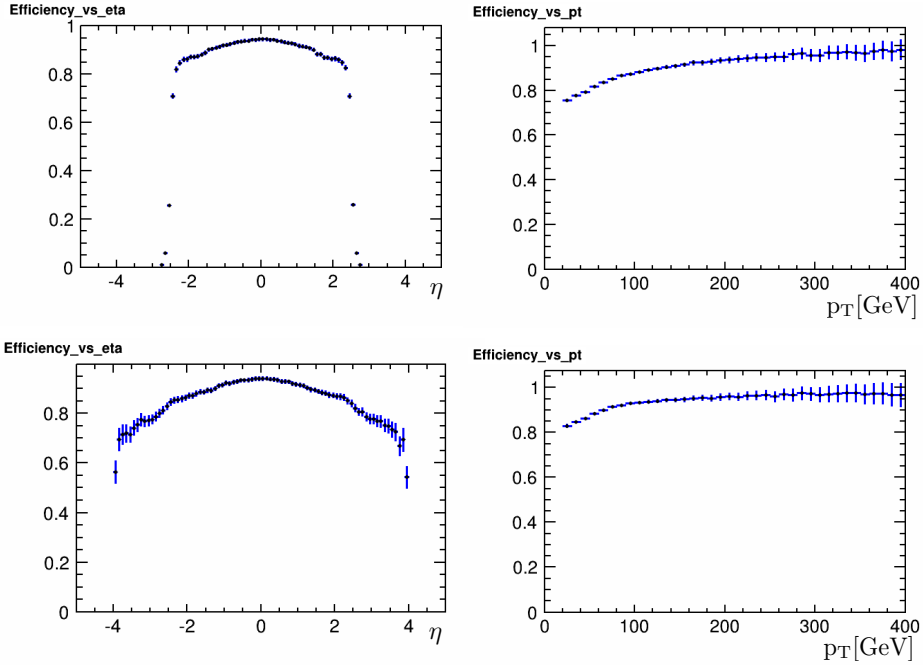

Figure 8.6: Muon reconstruction efficiency as a function of η or p_T. The Phase-I Not Aged scenario is considered on the top, the Phase-II one on the bottom.

section.

The selections applied at the event generation are listed in table 8.2.

sample	generator	cross section (fb)	event num.
EWK WZ SF H126	Phantom	3.924 ± 0.003	$2{\cdot}10^6$
EWK WZ SF noH	Phantom	4.022 ± 0.003	$2{\cdot}10^6$
EWK WZ DF H126	Phantom	3.930 ± 0.003	$2{\cdot}10^6$
EWK WZ DF noH	Phantom	4.033 ± 0.003	$2{\cdot}10^6$
QCD WZ SF	Phantom	16.26 ± 0.011	$2{\cdot}10^6$
QCD WZ DF	Phantom	16.31 ± 0.011	$2{\cdot}10^6$
QCD ZZ SF	Phantom	0.6280 ± 0.0004	$2{\cdot}10^6$
QCD ZZ DF	Phantom	0.6310 ± 0.0004	$2{\cdot}10^6$
semil. $t\bar{t}$	Powheg	$(242.5 \pm 0.2)10^3$	$16{\cdot}10^6$
W+3jets	Madgraph	$6.34 \cdot 10^3$	
DY+2jets	Madgraph	$5.5 \cdot 10^3$	
Z+3jets	Madgraph		$4{\cdot}10^6$
EWK WZ polarised	Madgraph		

Table 8.1: The list of generated samples, together with the generator used, the total number of produced events and the corresponding cross-section. Same Flavour-SF (Different Flavour-DF) corresponds to the events of the same(different) flavour leptons. H126 samples are the SM events while noH correspond to events without a Higgs boson.

selection	limit
minimum jet p_T	20 GeV
maximum jet η	6.5
minimum jet energy	20 GeV
minimum lepton energy	20 GeV
maximum lepton η	4.0
minimum 4-lepton invariant mass	130 GeV
minimum di-jet invariant mass	300 GeV
minimum di-jet rapidity difference	2
minimum di-lepton invariant mass	4 GeV

Table 8.2: The selections applied at event generation.

8.3.3 WZ Scattering Event Selection

The WZ scattering is studied in the three-lepton channel accompanied with two tagged jets where the Z boson decays into two leptons and the W boson into a lepton and a neutrino. To select WZ events, the existence of exactly three leptons, with p_{T} greater than 20 GeV, in the event is required. Two of the leptons should have same-flavour and opposite-charge to reconstruct the Z boson. If there are two flavours in the event, then the lepton which is not paired to form the Z boson is assumed to come from the W boson. For the cases, when all three leptons have the same flavour, all combinations of opposite-charge pairs are evaluated to reconstruct the di-lepton invariant mass and the combination that has a value closest to the Z boson invariant mass within $m_Z \pm 10$ GeV is taken and the unpaired lepton is assumed to come from the W boson. Events with additional loose leptons are vetoed to reduce the ZZ production contamination, where the loose and tight lepton definitions are defined in Sec. 8.3.2. The undetected neutrino momentum appears as missing energy in the detector and a cut of 30 GeV on missing transverse energy is applied to ensure the existence of the W boson in the event. In order to select deflected initial quarks, at least two jets with p_{T} larger than 30 GeV have to be detected and the first two highest p_{T} ones are identified as the tag jets coming from the scattering events. These selections applied up to here target at selecting WZ events with additional jets which contains of events of WZ EWK production $\mathrm{O}(\alpha_{ew}^6)$ and also WZ QCD production $\mathrm{O}(\alpha_{ew}^4\alpha_s^2)$. To suppress the QCD background and enhance the WZ scattering events among the EWK production the following additional cuts, specific to select the scattering events, are applied. The pseudorapidity separation of the tagged jets ($\Delta\eta_{jj}$) has to be larger than 4 units and the invariant mass of these two tag jets (m_{jj}) has to be larger than 600 GeV. The event selection cuts are optimised for the highest significance of the signal events using the ratio of $\mathrm{S}/\sqrt{S+B}$ where S and B are the expected signal and background yields passing the event selections respectively.

In this analysis, there are three stages, for which the following are calculated:

- the feasibility of the EWK production of WZ boson pairs with two jets *" inclusive analysis"*;
- the discovery of the longitudinal WZ scattering *" polarization analysis"*;
- the exclusion of a no-Higgs boson scenario *" H vs NoH analysis"*.

The selections common to the three analyses, explained above, are listed in table 8.3.

The expected signal and background distributions of the variables related with the VBS topology, $\mathrm{m}_{jj}, \Delta\eta_{jj}$ are shown in Fig. 8.7 for the Phase-II detector configuration with an integrated luminosity of 3 ab^{-1}. At the top plots

Selection	Limit
Tight Leptons	
minimum lepton p_T	20 GeV
maximum electron or muon relative isolation	0.25 (Phase-I 50 PU) 0.65 (Phase-II 140 PU)
Loose Leptons	
minimum lepton p_T	20 GeV
maximum electron or muon relative isolation	0.35 (Phase-I 50 PU) 0.75 (Phase-II 140 PU)
Jets	
minimum jet p_T	30 GeV
maximum jet η	4.7
cleaning jets from tight leptons	within $\Delta R = 0.3$
WZ Selections	
number of tight leptons	3
charge of the tight leptons	1 SFOC lepton pair, 1 charged lepton
number of extra leptons	0
number of jets	≥ 2
minimum MET value	30 GeV
Z mass cut for one SFOC lepton pair	$\lvert m_{ll} - m_Z \rvert < 10$ GeV
VBS Selections	
minimum tag jet m_{jj}	600 GeV
minimum tag jet $\Delta\eta_{jj}$	4

Table 8.3: The selections applied for to enhance scattering events of WZ EWK production with 2 jets

the VBS selections explained in Table 8.3 are not applied while at the bottom plots they are applied. After the VBS selections, the main remaining background is the WZ QCD production, shown in yellow, while the contribution of the ZZ events are very minor. In these plots, the blue continuous line corresponds to the EWK WZ scattering in the Standard Model, the dashed pink line corresponds to the EWK WZ scattering in absence of the Higgs boson, while the continuous red line corresponds to the difference between the two. Some other kinematical distributions with all the event selections applied are shown in Fig. 8.8.

8.4 Systematic Uncertainties

The systematic uncertainties affecting the WZ scattering analysis are listed in Tab. 8.4 together with their magnitudes, for each of the detector scenarios considered. The systematic uncertainties estimated for Phase I and Phase II detector scenarios are assumed to be identical, as the object reconstruction performances are expected to be similar [4]. For the aged-detector case, systematic uncertainties related with the object reconstruction are downgraded according to studies done by comparisons between the Phase II and Phase I

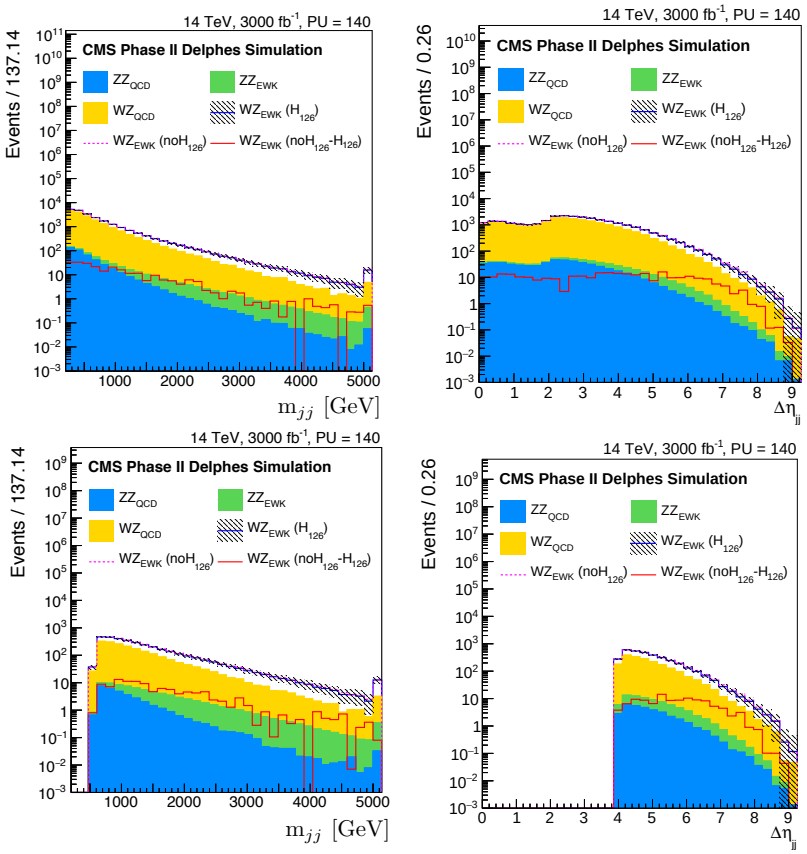

Figure 8.7: The m_{jj} and $\Delta\eta_{jj}$ of the two tagged jets are shown. The top plots show the distributions without the VBS cuts applied, while the bottom plots show the same distributions after the VBS cuts applied.

aged full simulation based on GEANT in [4].

As discussed in detail in Sec. 7.2, in order to estimate the experimental systematic uncertainties comparison with the data is needed. In the case of this study where the studies only rely on simulations, the experimental uncertainties are inherited from Run-I studies and treated in the same way as in Sec. 7.2. The theoretical uncertainties of the PDF and QCD scale are estimated in the same way as it is explained in Sec. 7.2. An acceptance uncertainty of 2% on the signal normalisation is also considered. Finally, also the statistical uncertainties of the simulation samples are taken into account as systematic uncertainty.

8.5 WZ Scattering Results

The expected analysis performances for the each detector scenarios are evaluated with the likelihood fits of the most sensitive variables for all the

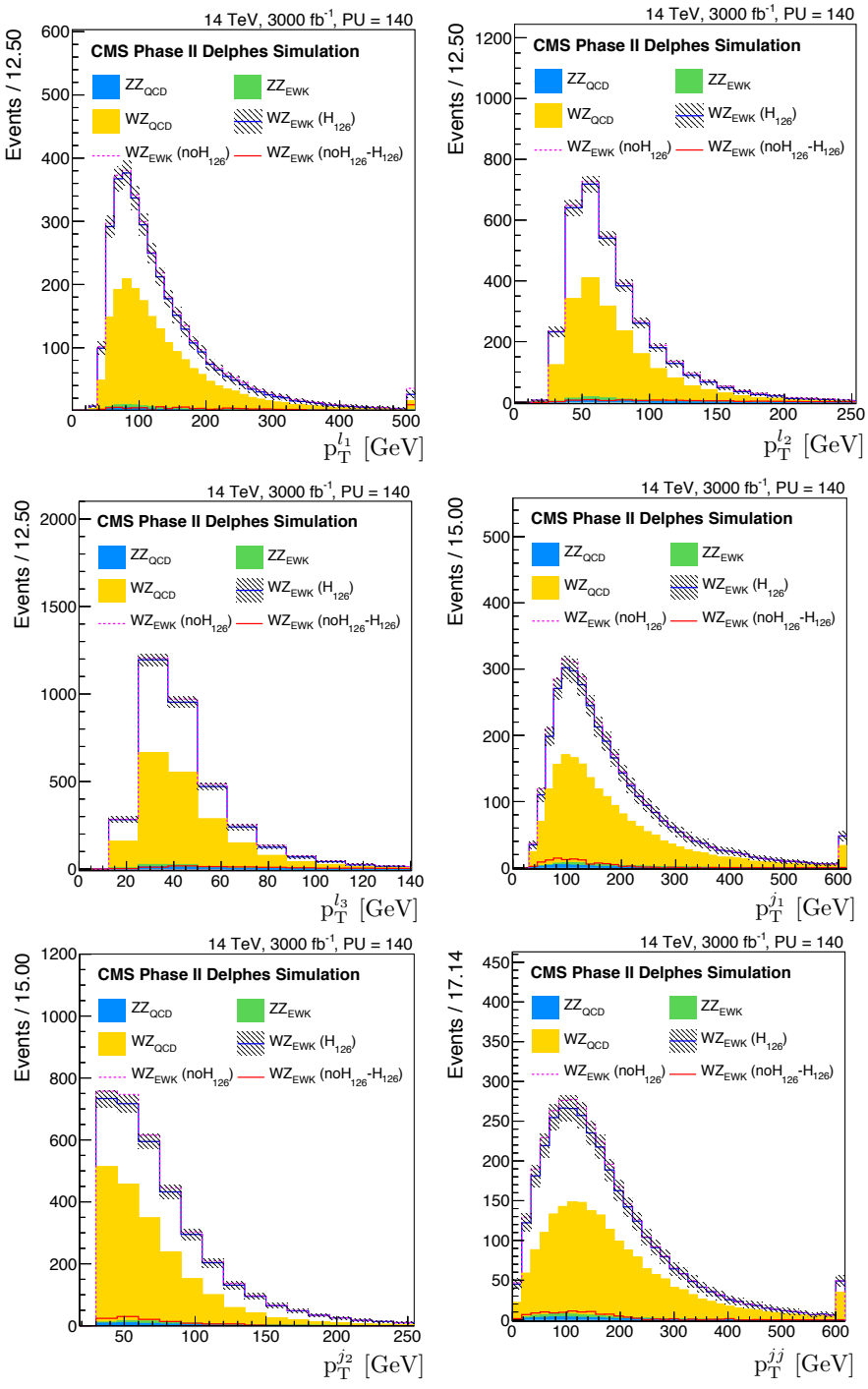

Figure 8.8: The transverse momenta of the leading, sub-leading and trailing leptons ($p_T^{l_1}$, $p_T^{l_2}$, $p_T^{l_3}$), two tagged jets ($p_T^{j_1}$, $p_T^{j_2}$) and of the tagged jet system (p_T^{jj}).

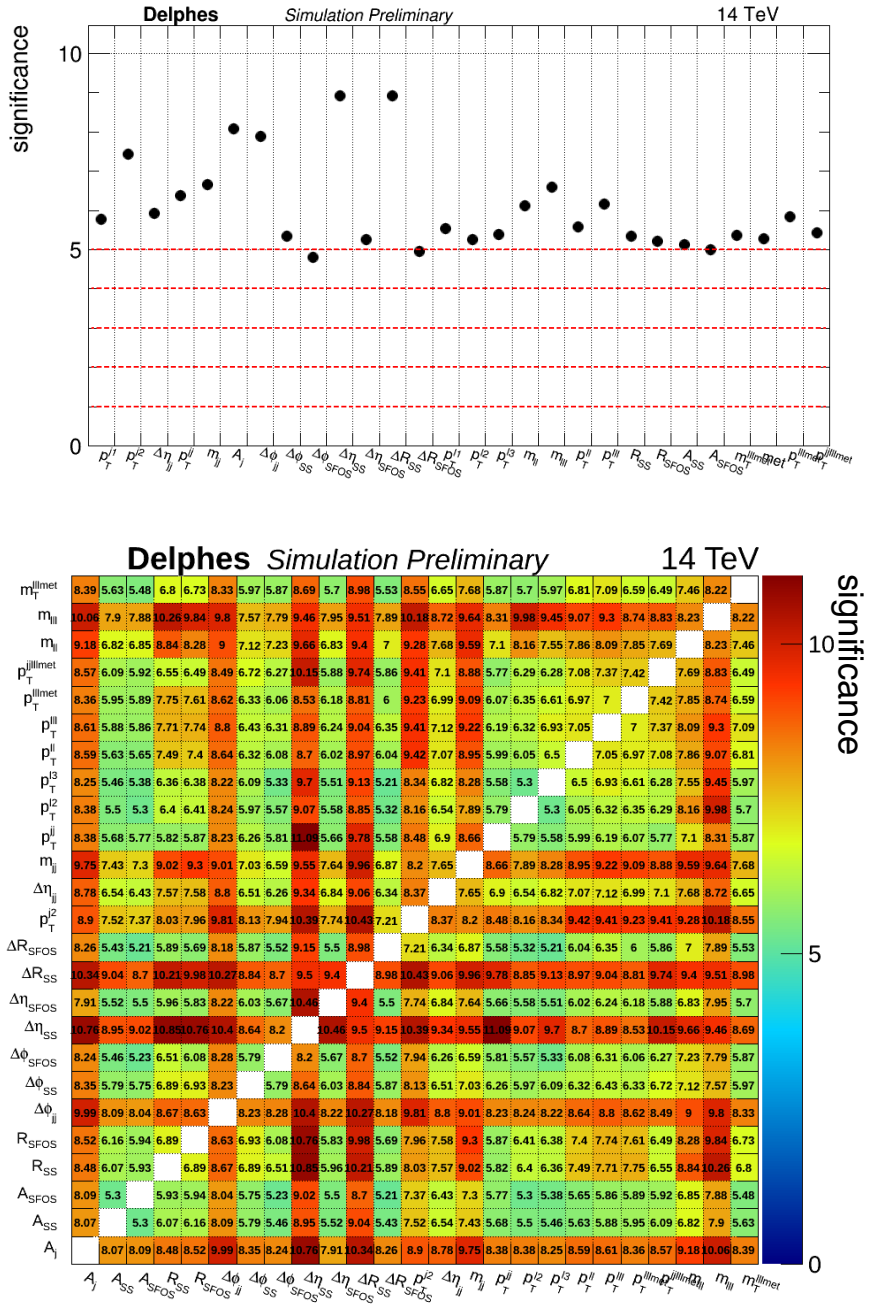

Figure 8.9: The 1D and 2D likelihood scans of the inclusive WZ EWK cross section measurement for the Phase I detector configuration for 3 ab^{-1} of integrated luminosity.

Systematic sources	Detector scenarios		
	Phase I	Phase I aged	Phase II
jet energy scale	1–3%	1.5–4%	1–3%
jet energy resol.	5%	6.5%	5%
muon energy scale	1%	2%	1%
muon energy resol.	1%	2%	1%
electron energy scale	2%	4%	2%
electron energy resol.	2%	4%	2%
lepton efficiency	2%	2%	2%
b-tag efficiency	4%	5.5%	4%
signal acceptance	2%	2%	2%
QCD scale choice	3%	3%	3%
PDF	7%	7%	7%
LHC luminosity	2.6%	2.6%	2.6%

Table 8.4: The systematics used for the WZ scattering analysis for each of the detector scenarios considered. The second and third columns indicate which systematics are used by the two analyses.

analysis benchmarks with the use of the statistical procedures explained in Sec. 7.3. The systematic uncertainties are treated in the same way as in $t\bar{t}Z$ analysis. Both one-dimensional(1D) distributions and all the possible two-dimensional(2D) distributions of the variables given in Tab. 8.5 are used for the fit. An example of the 1D and 2D likelihood scans of the inclusive WZ EWK cross section measurement for the Phase I detector configuration are shown in Fig. 8.9. For the final results 2D fits are used where a better sensitivity is obtained with respect to the 1D fit.

8.5.1 EWK WZ Scattering Result

The inclusive EWK cross section of the WZ boson production in association with two jets is determined by fitting the 2D distribution of $p_{\mathrm{T}}^{jj}, \Delta\eta_{ll}^{SS}$, where p_{T}^{jj} is the transverse momentum of the tag jets pair and $\Delta\eta_{ll}^{SS}$ is pseudorapidity separation between the two leptons of the same sign (SS). The distributions of these variables are shown in Fig. 8.8.

Given the fact that by the time of the HL-LHC the WZ scattering is expected to be discovered, that it is expected with 200 fb^{-1} of data, the important parameter to compare the detector performances is the precision of the cross section measurement. The expected precisions of the WZ EWK cross section for each detector scenario is shown in Fig. 8.11. Please note that the expected integrated luminosity at the end of HL-LHC is 3-4 ab^{-1}, while in

description	definition
η separation between tag jets	$\Delta\eta_{jj}$
φ separation between tag jets	$\Delta\varphi_{jj}$
tag jets invariant mass	m_{jj}
p_T of the tag jets	p_T^{jj}
p_T of the leading jet	$p_T^{j_1}$
p_T of the trailing jet	$p_T^{j_2}$
hadronic asymmetry	$A_j = \frac{p_T^{j_1} - p_T^{j_2}}{p_T^{j_1} + p_T^{j_2}}$
p_T of the leading lepton	$p_T^{\ell_1}$
p_T of the second leading lepton	$p_T^{\ell_2}$
p_T of the trailing lepton	$p_T^{\ell_3}$
φ separation between SF OS leptons (Z)	$\Delta\varphi(\ell,\ell)$
φ separation between SS leptons	$\Delta\varphi(\ell_1,\ell_2)$
η separation between Z leptons	$\Delta\eta_{ll}^{Z}$
η separation between SS leptons	$\Delta\eta_{ll}^{SS}$
ΔR between Z leptons	ΔR_{ll}^{Z}
ΔR between SS leptons	ΔR_{jj}^{SS}
leptonic asymmetry of SS	$A_\ell = \frac{p_T^{\ell_1} - p_T^{\ell_2}}{p_T^{\ell_1} + p_T^{\ell_2}}$
leptonic asymmetry of SFOS	$A_\ell = \frac{p_T^{\ell_1} - p_T^{\ell_2}}{p_T^{\ell_1} + p_T^{\ell_2}}$
transverse missing energy	MET
R-variable of SS leptons	$R_{SS} = \frac{p_T^{\ell_1} \cdot p_T^{\ell_2}}{p_T^{j_1} \cdot p_T^{j_2}}$
R-variable of SFOS leptons	$R_Z = \frac{p_T^{\ell_1} \cdot p_T^{\ell_2}}{p_T^{j_1} \cdot p_T^{j_2}}$
invariant mass of the SFOS leptons	m_{ll}
invariant mass of three leptons	m_{lll}
transverse momentum of SFOS leptons	p_T^{ll}
transverse momentum of three leptons	p_T^{lll}
transverse momentum of three leptons and MET	$p_T^{lll,met}$
transverse momentum of three leptons, two jets and MET	$p_T^{lll,jj,met}$

Table 8.5: The variables investigated in order to obtain the best signal sensitivity for the WZ scattering study.

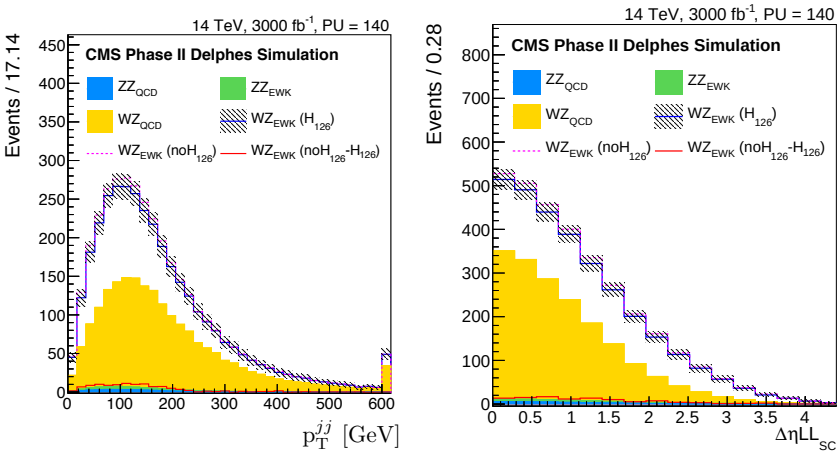

Figure 8.10: Distributions of the p_T of the di-jet pair and the difference in pseudorapidity between the same sign leptons for the Phase II detector, after the WZ selection. The signal and the backgrounds are normalised to the integrated luminosity of 3 ab^{-1} . In particular, the blue continuous line corresponds to the EWK di-boson scattering in the Standard Model, the dashed pink one corresponds to the EWK di-boson scattering in absence of the Higgs boson, while the continuous red line corresponds to the difference between the two.

the representation of the plots a range up to 6 ab^{-1} is shown. This is to see if there is a gain on the expected sensitivity with the collection of more data. The expected uncertainty of the cross section with 3 ab^{-1} of data, uncertainties of 9% can be achieved with the Phase I and Phase II detector scenarios, while the Phase I aged detector would provide uncertainties of the order of 12%. Therefore the upgraded Phase II detector in high luminosity and high PU environment is expected perform as good as the Phase I detector with 50 PU. Considering the fact that at the HL-LHC the expected total integrated is 3-4 ab^{-1}, it is possible to improve the precision of this measurement with more data.

8.5.2 Sensitivity to the Longitudinal WZ Scattering Component

The total vector boson scattering can be composed into three components depending on the polarization of the final-state vector bosons. These are: both of the bosons being longitudinally polarised (LL), both of the bosons being transversely polarised (TT) and the mixed case (LT). To determine the sensitivity of the longitudinal component of the cross section, the LL component is treated as the signal while TT and LT components are treated together with the main backgrounds. The best variable pair for the longitudinal scattering is found to be $\Delta\varphi^{jj}, p_T^{l1}$. Fig. 8.12 shows the distributions of these two variables for the LL, TT and TL components of the EWK WZ scattering together with the backgrounds.

The performance obtained in this configuration, is shown in Fig. 8.13 as a

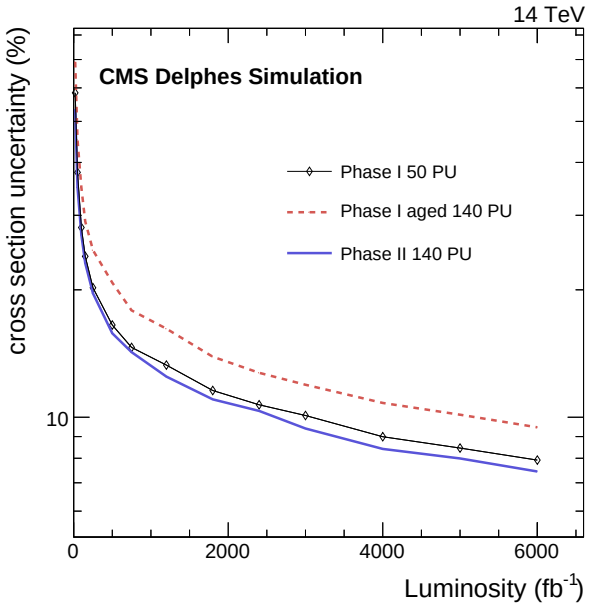

Figure 8.11: The expected total uncertainty for the various detector scenarios for the EWK WZ cross section measurement as a function of the collected luminosity.

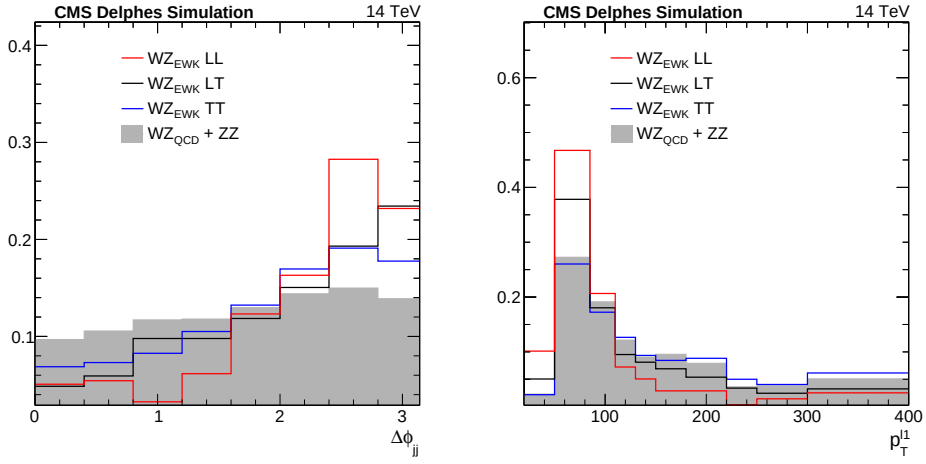

Figure 8.12: Left: the azimuthal angle difference distribution between the two tag jets for different WZ polarization states. Right: the distribution of the highest p_T lepton in the event for different WZ polarization states. Both histograms are normalised to the unity.

function of the collected luminosity for the three detector scenarios. For the Phase II detector with 3 ab^{-1} of data, an expected significance for the LL component of 1.4 sigma can be reached while values lower by $\sim$ 25% for the Phase I aged detector are expected. Even with the large amount of data collected in HL-LHC and with larger amount of data (up to 6 ab^{-1}) this measurement remains challenging.

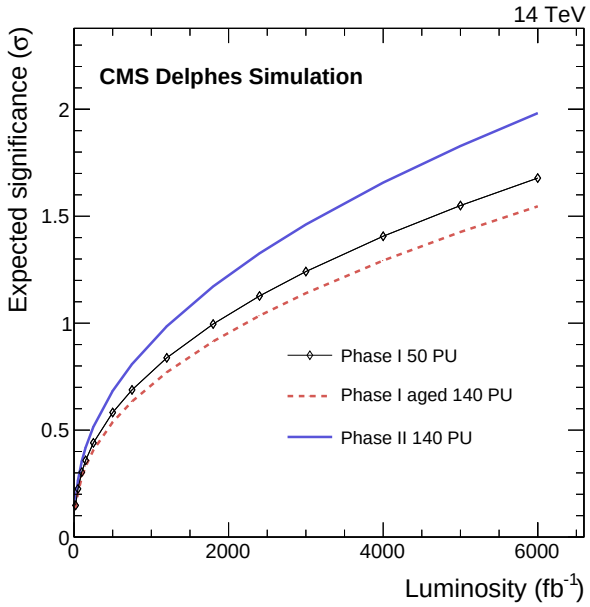

Figure 8.13: The expected discovery significance for the longitudinal vector boson scattering for the various detector scenarios as a function of the collected luminosity for the WZ analysis.

8.5.3 Partial Unitarisation Scenario

As explained in Chapter 2, the Higgs boson unitarizes the cross section of the VBS process. Therefore the measurement of the VBS cross section can be used for testing this in a model independent way. This can be achieved by comparing the SM case with the case where the Higgs boson does not participate in the unitarization called as no-Higgs (noH) scenario. The analysis selections are applied in the same way to SM and noH cases where an enhancement of the cross section is expected for the noH case. The SM expected distribution is shown in blue line in Fig. 8.14, while the no-Higgs case is shown in pink line in the same plot. In order to determine the separation power of these two cases, the difference between the SM and the no-Higgs case is treated as an artificial signal while the SM WZ scattering events are considered as part of the background. This is shown by the red line in Fig. 8.14. Therefore to asses the sensitivity to noH scenario, the expected 95% CL exclusion limits are set in terms of the strength modifier (μ) of the noH $-$ H hypothesis.

The highest expected exclusion power for the no-Higgs scenario with respect to the Standard Model is obtained by fitting the 2D templates of $p_T^{l3}, \Delta\eta_{jj}$. The 95% exclusion limit as a function of the total integrated luminosity is shown in Fig. 8.15. If the 95% confidence level exclusion is reached for $\mu = 1$, the analysis would be sensitive to the case where the Higgs boson does not participate to the WZ scattering unitarization, while for values

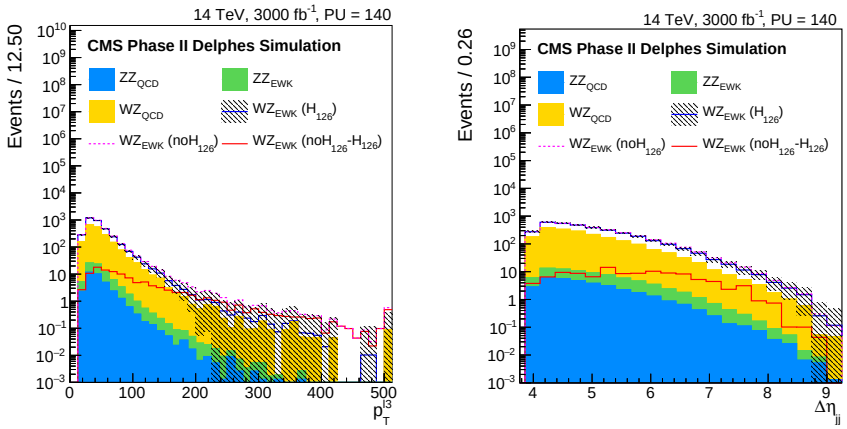

Figure 8.14: Distributions of the p_T of the trailing lepton and the difference in pseudorapidity between tag jets for the Phase II detector, after the WZ selection. The signal and the backgrounds are normalised to the integrated luminosity of 3 ab^{-1} . In particular, the blue continuous line corresponds to the EWK di-boson scattering in the Standard Model, the dashed pink one corresponds to the EWK di-boson scattering in absence of the Higgs boson, while the continuous red line corresponds to the difference between the two.

of μ smaller than one, the analysis will be sensitive to scenarios of partial unitarization. The exclusion limit is expected to be 1 with a dataset of 600 fb^{-1} and with 3 ab^{-1} the exclusion limit is expected to be around 0.35 with Phase II detector configuration. As can be seen from Fig. 8.15, the upgraded CMS detector is more sensitive than the aged version of the current one and recovers the same performances one would get with the same luminosity, acquired during data conditions of the LHC Run-I.

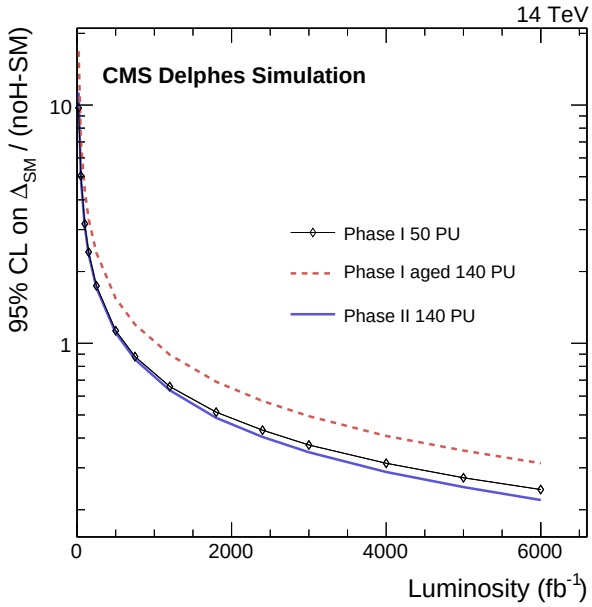

Figure 8.15: The expected 95% CL exclusion power for the no-Higgs scenario as a function of the integrated luminosity for the WZ analysis. The limit is expressed as deviation from the Standard Model divided by the difference of the no-Higgs case from the Standard Model itself.

Part III

Conclusions and Outlook

Chapter 9

Conclusions and Outlook

The LHC has been running since 2010 at centre of mass energies of 7-8 and 13 TeV. With the discovery of the Higgs boson all the particles predicted by the SM are discovered. So far the LHC data shows very good agreement with the predictions of the SM. Yet, there are puzzles that can not be answered by the SM which motivates beyond the SM. Therefore this brings the expectation of new discoveries in the near future, as well, motivates the extended projects of LHC such as HL-LHC or new experiments such as linear electron colliders. On the other hand, the precision measurements are crucial to refine the Standard Model and test its overall consistency.

The first analysis presented in this thesis is the cross section measurement of the top-antitop pair production in association with a Z boson at a center-of-mass energy of 13 TeV at the LHC. The data sample corresponds to an integrated luminosity of 35.9 fb^{-1}, collected by the CMS experiment in 2016. I performed this analysis using pp collision events with three leptons. Such events are expected to provide best sensitivity to $t\bar{t}Z$ production as they lead to relatively small background. The production cross section of $t\bar{t}Z$ using such events is found to be $1.01^{+0.09}_{-0.09}(\text{stat.})^{+0.10}_{-0.11}(\text{sys.})\text{pb}$. Additionally, I combined the result presented in this thesis with $t\bar{t}Z$ production in four-lepton final state, presented in [8]. The combined result of the three and four-lepton final states of the $t\bar{t}Z$ production is found to be $0.99^{+0.09}_{-0.08}(\text{stat.})^{+0.12}_{-0.10}(\text{sys.})\text{pb}$. This result is the most precise measurement of this process up to date.

The uncertainty of the cross section measurement of the $t\bar{t}Z$ production is dominated by the systematic uncertainty, meaning that the gain on the precision of this measurement will not improve significantly with more available data, unless the systematic uncertainties are reduced. Therefore, in order to make this measurement more precise it is essential to work on methods to reduce the systematic uncertainties. The sources of the systematic uncertainties which lead to the largest contributions are imprecise knowledge on lepton identification and trigger efficiencies, $t\bar{t}X$ background and luminos-

ity. The current uncertainties assigned to lepton identification has room for improvement in the following way. As explained in Sec. 6.6.2, the standard Tag-and-Probe method uses DY events in data and simulation to validate lepton efficiency in simulation and assess the related uncertainty. However, DY events typically lead to leptons with large p_{T}, such that for low p_{T} leptons few statistics is available for applying the Tag-And-Probe method. Therefore, the larger systematics assigned to low p_{T} muons can be improved by applying the Tag-and-Probe tool within the invariant mass range of J/Ψ. This will withdraw the additional uncertainty used for low p_{T} muons. In case of trigger uncertainty, a conservative uncertainty was assigned. This can be improved by reworking the method in more detail and possibly validating the measured efficiencies in orthogonal data samples. In addition, with the larger datasets available, the transverse momentum thresholds assigned to lepton triggers will need to be increased in order to cope with processing high event rates. This would cause a decrease in the trigger efficiency, in particular when the leading lepton p_{T} is lower than the trigger threshold. In this case, the combination of single-, double- and trilepton triggers can be used to recover this in-efficiency and it may potentially lead to lower systematic uncertainties. The other dominant systematic uncertainty stems from the imprecise knowledge on backgrounds grouped in $\mathrm{t\bar{t}}X$. These background processes, such as $\mathrm{t\bar{t}H}$, are expected to be measured more precisely with the fore-coming data. This will require large data volumes. However this is the only way to reduce the uncertainty of these backgrounds beyond the theoretical uncertainties. Additionally, the differential measurements of $\mathrm{t\bar{t}}X$ will help improving the modelling. Therefore, the assigned uncertainty to these backgrounds will be smaller which is expected to, as well, reduce the impact of this uncertainty on this measurement.

The result presented in this thesis in combination with the result of $\mathrm{t\bar{t}W}$ measurement, as well as the expected $\mathrm{t\bar{t}H}$ contribution, is used to search for new physics using the effective field theory approach [8]. The SM Lagrangian is extended in terms of the dimension six operators where the Wilson coefficient of these operators parameterise the strength of new physics interactions. The CL limits to eight operators have been put which are of particular interest because they change the expected cross sections of $\mathrm{t\bar{t}Z}$, $\mathrm{t\bar{t}W}$, or $\mathrm{t\bar{t}H}$.

Considering the future prospects of the $\mathrm{t\bar{t}Z}$, another test of new physics can be performed by measuring the top quark and Z boson coupling. These constants were evaluated at 7 TeV [83] and can be extrapolated to any energy linearly which was done at 8 TeV[14]. The increased data sets will allow constraining the current limits and also will enable to study the differential distributions of $\mathrm{t\bar{t}Z}$. Additionally, cross section ratios of $\mathrm{t\bar{t}Z}/\mathrm{t\bar{t}H}$ and $\mathrm{t\bar{t}Z}/\mathrm{t\bar{t}}\gamma$ will be accessible. Such studies will have the advantage of the cancellation of the common uncertainties, such as luminosity, trigger, etc.

The second analysis summarised in this thesis is the study of WZ scattering in view of the HL-LHC. This study was performed in order to estimate the performance of the upgraded CMS detector for the HL-LHC. The results are presented by comparing the sensitivity of the future CMS detector with the current one and with the detector which will be degraded due to radiation. It is shown that the upgraded detector will recover the performance of the current detector for the precision of the WZ cross section measurement and for putting limits on the partial unitarization. While in case of the expected significance of the longitudinal component of the WZ scattering production, the upgraded detector will perform better than the current detector up to 25%. The WZ scattering, still being an un-observed process, is expected to be discovered with 200 fb^{-1} of data. This amount of data is expected to be delivered by the end of 2021 by the LHC.

The longitudinal component of the WZ scattering has an expected significance of 1.5 (2)σ with 3 (6)ab^{-1} of data. It needs to be taken into account that, these results are the expectations of a first and conservative study and by the time of the HL-LHC there will be improvements that can increase the sensitivity. Nevertheless, to study the longitudinal component of the WZ scattering in HL-LHC timeline still remains very challenging. However, the combination with other vector boson scattering states will increase the sensitivity of the longitudinal component of vector boson scattering. The WW same sign scattering in view of the HL-LHC, for the same upgrade and detector configurations described in this thesis, is studied and reported in [4], [5]. Additionally, the combined results of the WW same sign and WZ scattering are also reported in the same references. The expected sensitivity to the longitudinal component of WW same-sign and WZ scattering is reported as 2.75σ. This result may be marginally improved with the combination of less sensitive vector boson scattering states WW opposite-sign and ZZ scattering.

In terms of the detector upgrades of the HL-LHC, the tracker extension to $\eta = 4$ has an important impact on this analysis. The outgoing jets are expected to be forward so that with this extension the signal jets are reconstructed up to $\eta = 4$. It is also possible to constrain the structure of quartic vector boson interactions in the framework of dimension-eight effective field theory operators via studying the vector boson scattering. These constraints were studied at 8 TeV, while there are no results on these at 13 TeV [7], [128]. This will be one of the interesting studies to be performed with more available data.

Bibliography

[1] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321 (cit. on pp. 1, 11).

[2] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508 (cit. on pp. 1, 11).

[3] *ECFA-CERN Workshop on Large Hadron Collider in the LEP tunnel*. CERN. Geneva: CERN, 1984. URL: http://cds.cern.ch/record/154938 (cit. on p. 1).

[4] D Contardo et al. *Technical Proposal for the Phase-II Upgrade of the CMS Detector*. Tech. rep. CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02. Geneva, June 2015. URL: https://cds.cern.ch/record/2020886 (cit. on pp. 2, 119–121, 124, 129, 130, 145).

[5] *Prospects for the study of vector boson scattering in same sign WW and WZ interactions at the HL-LHC with the upgraded CMS detector*. Tech. rep. CMS-PAS-SMP-14-008. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2220831 (cit. on pp. 2, 145).

[6] Vardan Khachatryan et al. "Study of vector boson scattering and search for new physics in events with two same-sign leptons and two jets". In: *Phys. Rev. Lett.* 114.5 (2015), p. 051801. DOI: 10.1103/PhysRevLett.114.051801. arXiv: 1410.6315 [hep-ex] (cit. on p. 2).

[7] Georges Aad et al. "Measurements of $W^{\pm}Z$ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector and limits on anomalous gauge boson self-couplings". In: *Phys. Rev.* D93.9 (2016), p. 092004. DOI: 10.1103/PhysRevD.93.092004. arXiv: 1603.02151 [hep-ex] (cit. on pp. 2, 145).

[8] Albert M Sirunyan et al. "Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: (2017). arXiv: 1711.02547 [hep-ex] (cit. on pp. 3, 112, 113, 143, 144).

[9] *Measurement of the cross section of top quark pair production in association with a Z boson in pp collisions at 13 TeV*. Tech. rep. CMS-PAS-TOP-16-009. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2139263 (cit. on p. 3).

[10] *Measurement of the top pair-production in association with a W or Z boson in pp collisions at 13 TeV*. Tech. rep. CMS-PAS-TOP-16-017. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2205283 (cit. on p. 3).

[11] Serguei Chatrchyan et al. "Measurement of associated production of vector bosons and top quark-antiquark pairs at sqrt(s) = 7 TeV". In: *Phys. Rev. Lett.* 110 (2013), p. 172002. DOI: 10.1103/PhysRevLett.110.172002. arXiv: 1303.3239 [hep-ex] (cit. on p. 3).

[12] Vardan Khachatryan et al. "Measurement of top quark-antiquark pair production in association with a W or Z boson in pp collisions at $\sqrt{s} = 8$ TeV". In: *Eur. Phys. J.* C74.9 (2014), p. 3060. DOI: 10.1140/epjc/s10052-014-3060-7. arXiv: 1406.7830 [hep-ex] (cit. on p. 3).

[13] Georges Aad et al. "Measurement of the $t\bar{t}W$ and $t\bar{t}Z$ production cross sections in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 11 (2015), p. 172. DOI: 10.1007/JHEP11(2015)172. arXiv: 1509.05276 [hep-ex] (cit. on p. 3).

[14] Vardan Khachatryan et al. "Observation of top quark pairs produced in association with a vector boson in pp collisions at $\sqrt{s} = 8$ TeV". In: *JHEP* 01 (2016), p. 096. DOI: 10.1007/JHEP01(2016)096. arXiv: 1510.01131 [hep-ex] (cit. on pp. 3, 22, 144).

[15] Deniz Poyraz. "Vector Boson Scattering prospects for High-Luminosity LHC at CMS in the WZ final state". In: *PoS* LeptonPhoton2015 (2016), p. 112 (cit. on p. 4).

[16] Deniz Poyraz. "Measurement of the top quark properties in the production and decays of top anti-top pair events at CMS". In: *PoS* DIS2016 (2016), p. 146. arXiv: 1608.05640 [hep-ex] (cit. on p. 4).

[17] Andrew Purcell. "Go on a particle quest at the first CERN webfest. Le premier webfest du CERN se lance la conqute des particules". In: BUL-NA-2012-269. 35/2012 (Aug. 2012), p. 10. URL: https://cds.cern.ch/record/1473657 (cit. on p. 7).

[18] David J Griffiths. *Introduction to elementary particles; 2nd rev. version*. Physics textbook. New York, NY: Wiley, 2008. URL: https://cds.cern.ch/record/111880 (cit. on p. 6).

[19] F. Halzen and Alan D. Martin. *QUARKS AND LEPTONS: AN INTRODUCTORY COURSE IN MODERN PARTICLE PHYSICS*. 1984. ISBN: 0471887412, 9780471887416 (cit. on p. 9).

[20] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264 (cit. on p. 10).

[21] the SLD Electroweak. "A Combination of preliminary electroweak measurements and constraints on the standard model". In: (2003). arXiv: hep-ex/0312023 [hep-ex] (cit. on p. 14).

[22] Georges Aad et al. "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV". In: *JHEP* 08 (2016), p. 045. DOI: 10.1007/JHEP08(2016)045. arXiv: 1606.02266 [hep-ex] (cit. on p. 16).

[23] Y. Fukuda et al. "Evidence for oscillation of atmospheric neutrinos". In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. arXiv: hep-ex/9807003 [hep-ex] (cit. on p. 16).

[24] Q. R. Ahmad et al. "Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory". In: *Phys. Rev. Lett.* 89 (2002), p. 011301. DOI: 10.1103/PhysRevLett.89.011301. arXiv: nucl-ex/0204008 [nucl-ex] (cit. on p. 16).

[25] N. Jarosik et al. "Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results". In: 192, 14 (Feb. 2011), p. 14. DOI: 10.1088/0067-0049/192/2/14. arXiv: 1001.4744 (cit. on p. 17).

[26] Felix Kahlhoefer. "Review of LHC Dark Matter Searches". In: *Int. J. Mod. Phys.* A32.13 (2017), p. 1730006. DOI: 10.1142/S0217751X1730006X. arXiv: 1702.02430 [hep-ph] (cit. on p. 17).

[27] P. J. E. Peebles and Bharat Ratra. "The Cosmological constant and dark energy". In: *Rev. Mod. Phys.* 75 (2003), pp. 559–606. DOI: 10.1103/RevModPhys.75.559. arXiv: astro-ph/0207347 [astro-ph] (cit. on p. 17).

[28] A. D. Sakharov. "Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe". In: *Pisma Zh. Eksp. Teor. Fiz.* 5 (1967). [Usp. Fiz. Nauk161,no.5,61(1991)], pp. 32–35 (cit. on p. 17).

[29] M Peskin. "The matter with antimatter". In: *Nature* 419 (2002), no pagination. URL: https://cds.cern.ch/record/717282 (cit. on p. 17).

[30] F. Abe et al. "Observation of top quark production in $\bar{p}p$ collisions". In: *Phys. Rev. Lett.* 74 (1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. arXiv: hep-ex/9503002 [hep-ex] (cit. on p. 18).

[31] S. Abachi et al. "Observation of the top quark". In: *Phys. Rev. Lett.* 74 (1995), pp. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632. arXiv: hep-ex/9503003 [hep-ex] (cit. on p. 18).

[32] K. A. Olive et al. "Review of Particle Physics". In: *Chin. Phys.* C38 (2014), p. 090001. DOI: 10.1088/1674-1137/38/9/090001 (cit. on pp. 18, 48, 107, 109).

[33] Frank-Peter Schilling. "Top Quark Physics at the LHC: A Review of the First Two Years". In: *Int. J. Mod. Phys.* A27 (2012), p. 1230016. DOI: 10.1142/S0217751X12300165. arXiv: 1206.4484 [hep-ex] (cit. on pp. 18, 19).

[34] C P Yuan. "Top Quark and Electroweak Symmetry Breaking Mechanism". In: hep-ph/9809536. MSU-HEP-80828 (Sept. 1998), 9 p. URL: http://cds.cern.ch/record/366378 (cit. on p. 18).

[35] Morad Aaboud et al. "Probing the W tb vertex structure in t-channel single-top-quark production and decay in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP* 04 (2017), p. 124. DOI: 10.1007/JHEP04(2017)124. arXiv: 1702.08309 [hep-ex] (cit. on p. 20).

[36] Vardan Khachatryan et al. "Measurement of top quark polarisation in t-channel single top quark production". In: *JHEP* 04 (2016), p. 073. DOI: 10.1007/JHEP04(2016)073. arXiv: 1511.02138 [hep-ex] (cit. on p. 20).

[37] Kenneth G. Wilson. "Nonlagrangian models of current algebra". In: *Phys. Rev.* 179 (1969), pp. 1499–1512. DOI: 10.1103/PhysRev.179.1499 (cit. on p. 22).

[38] Olga Bessidskaia Bylund et al. "Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD". In: *JHEP* 05 (2016), p. 052. DOI: 10.1007/JHEP05(2016)052. arXiv: 1601.08193 [hep-ph] (cit. on p. 22).

[39] Christoph Englert et al. "Higgs coupling measurements at the LHC". In: *Eur. Phys. J.* C76.7 (2016), p. 393. DOI: 10.1140/epjc/s10052-016-4227-1. arXiv: 1511.05170 [hep-ph] (cit. on p. 22).

[40] B. Grzadkowski et al. "Dimension-Six Terms in the Standard Model Lagrangian". In: *JHEP* 10 (2010), p. 085. DOI: 10.1007/JHEP10(2010)085. arXiv: 1008.4884 [hep-ph] (cit. on p. 22).

[41] Adam Alloul, Benjamin Fuks, and Vernica Sanz. "Phenomenology of the Higgs Effective Lagrangian via FEYNRULES". In: *JHEP* 04 (2014), p. 110. DOI: 10.1007/JHEP04(2014)110. arXiv: 1310.5150 [hep-ph] (cit. on p. 22).

[42] CMS Collaboration. "Measurement of top pair-production in association with a W or Z boson in pp collisions at 13 TeV". In: (2017) (cit. on p. 22).

[43] Ana Alboteanu, Wolfgang Kilian, and Juergen Reuter. "Resonances and Unitarity in Weak Boson Scattering at the LHC". In: *JHEP* 11 (2008), p. 010. DOI: 10.1088/1126-6708/2008/11/010. arXiv: 0806.4145 [hep-ph] (cit. on p. 23).

[44] John M. Cornwall, David N. Levin, and George Tiktopoulos. "Derivation of gauge invariance from high-energy unitarity bounds on the S matrix". In: *Phys. Rev. D* 10 (4 Aug. 1974), pp. 1145–1167. DOI: 10.1103/PhysRevD.10.1145. URL: https://link.aps.org/doi/10.1103/PhysRevD.10.1145 (cit. on p. 23).

[45] C. E. Vayonakis. "Born helicity amplitudes and cross-sections in non-Abelian gauge theories". In: *Lettere al Nuovo Cimento (1971-1985)* 17.11 (Nov. 1, 1976), pp. 383–387. ISSN: 1827-613X. DOI: 10.1007/BF02746538. URL: https://doi.org/10.1007/BF02746538 (cit. on p. 23).

[46] Michael S. Chanowitz and Mary K. Gaillard. "The TeV physics of strongly interacting W's and Z's". In: *Nuclear Physics B* 261.Supplement C (1985), pp. 379–431. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(85)90580-2. URL: http://www.sciencedirect.com/science/article/pii/0550321385905802 (cit. on p. 23).

[47] Benjamin W. Lee, C. Quigg, and H. B. Thacker. "Strength of Weak Interactions at Very High Energies and the Higgs Boson Mass". In: *Phys. Rev. Lett.* 38 (16 Apr. 1977), pp. 883–885. DOI: 10.1103/PhysRevLett.38.883. URL: https://link.aps.org/doi/10.1103/PhysRevLett.38.883 (cit. on p. 23).

[48] Benjamin W. Lee, C. Quigg, and H. B. Thacker. "Weak interactions at very high energies: The role of the Higgs-boson mass". In: *Phys. Rev. D* 16 (5 Sept. 1977), pp. 1519–1531. DOI: 10.1103/PhysRevD.16.1519. URL: https://link.aps.org/doi/10.1103/PhysRevD.16.1519 (cit. on p. 23).

[49] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex] (cit. on p. 23).

[50] Serguei Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex] (cit. on p. 23).

[51] Celine Degrande et al. "Effective Field Theory: A Modern Approach to Anomalous Couplings". In: *Annals Phys.* 335 (2013), pp. 21–32. DOI: 10.1016/j.aop.2013.04.016. arXiv: 1205.4231 [hep-ph] (cit. on p. 23).

[52] O. J. P. Eboli, M. C. Gonzalez-Garcia, and J. K. Mizukoshi. "p p — > j j e+- mu+- nu nu and j j e+- mu-+ nu nu at O(alpha(em)**6) and O(alpha(em)**4 alpha(s)**2) for the study of the quartic electroweak gauge boson vertex at CERN LHC". In: *Phys. Rev.* D74 (2006), p. 073005. DOI: 10.1103/PhysRevD.74.073005. arXiv: hep-ph/0606118 [hep-ph] (cit. on p. 23).

[53] Roberto Aloisio et al. "Selected Topics in Cosmic Ray Physics". In: (2017). arXiv: 1707.06147 [astro-ph.HE] (cit. on p. 27).

[54] "Underground Area 1 collaboration, UA1: Technical Notes: Data acquisition". URL: http://cds.cern.ch/record/87128 (cit. on p. 28).

[55] P. Landshoff. "The Large Electron Positron Collider (LEP) for Particle Physics". In: *Contemp. Phys.* 22 (1951), pp. 291–298. DOI: 10.1080/00107518108231532 (cit. on p. 28).

[56] Timo A. Aaltonen. "Combination of CDF and DO results on the mass of the top quark using up to 8.7 fb^{-1} at the Tevatron". In: (2013). arXiv: 1305.3929 [hep-ex] (cit. on p. 28).

[57] Oliver Sim Bruning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. URL: https://cds.cern.ch/record/782076 (cit. on pp. 28, 29).

[58] Werner Herr and B Muratori. "Concept of luminosity". In: (2006). URL: https://cds.cern.ch/record/941318 (cit. on p. 31).

[59] K. Aamodt et al. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002 (cit. on p. 31).

[60] G. Aad et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003 (cit. on p. 31).

[61] S. Chatrchyan et al. "The CMS Experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004 (cit. on p. 32).

[62] A. Augusto Alves Jr. et al. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005 (cit. on p. 32).

[63] *The CMS tracker: addendum to the Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 2000. URL: https://cds.cern.ch/record/490194 (cit. on p. 36).

[64] *The CMS tracker: addendum to the Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 2000. URL: https://cds.cern.ch/record/490194 (cit. on p. 36).

[65] The CMS Collaboration. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (2008), S08004. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08004 (cit. on p. 37).

[66] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/349375 (cit. on p. 37).

[67] Philippe Bloch et al. *Changes to CMS ECAL electronics: addendum to the Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 2002. URL: http://cds.cern.ch/record/581342 (cit. on p. 37).

[68] P. Adzic et al. "Energy resolution of the barrel of the CMS electromagnetic calorimeter". In: *JINST* 2 (2007), P04004. DOI: 10.1088/1748-0221/2/04/P04004 (cit. on p. 38).

[69] S. Abdullin et al. "The CMS barrel calorimeter response to particle beams from 2-GeV/c to 350-GeV/c". In: *Eur. Phys. J.* C60 (2009). [Erratum: Eur. Phys. J.C61,353(2009)], pp. 359–373. DOI: 10.1140/epjc/s10052-009-0959-5,10.1140/epjc/s10052-009-1024-0 (cit. on p. 39).

[70] *The CMS muon project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/343814 (cit. on p. 40).

[71] Serguei Chatrchyan et al. "The performance of the CMS muon detector in proton-proton collisions at sqrt(s) = 7 TeV at the LHC". In: *JINST* 8 (2013), P11002. DOI: 10.1088/1748-0221/8/11/P11002. arXiv: 1306.6905 [physics.ins-det] (cit. on p. 40).

[72] Stefan Hoche. "Introduction to parton-shower event generators". In: *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*. 2015, pp. 235–295. DOI: 10.1142/9789814678766_0005. arXiv: 1411.4085 [hep-ph]. URL: https://inspirehep.net/record/1328513/files/arXiv:1411.4085.pdf (cit. on p. 46).

[73] William A. Bardeen et al. "Deep-inelastic scattering beyond the leading order in asymptotically free gauge theories". In: *Phys. Rev. D* 18 (11 Dec. 1978), pp. 3998–4017. DOI: 10.1103/PhysRevD.18.3998. URL: https://link.aps.org/doi/10.1103/PhysRevD.18.3998 (cit. on p. 47).

[74] John Collins. *Foundations of Perturbative QCD*. 1st ed. Cambridge Universit Press, July 2011. ISBN: 978-0-521-85533-4 (cit. on p. 47).

[75] Richard Keith Ellis, William James Stirling, and Bryan R Webber. *QCD and collider physics*. Cambridge monographs on particle physics, nuclear physics, and cosmology. Photography by S. Vascotto. Cambridge: Cambridge University Press, 2003. URL: https://cds.cern.ch/record/318585 (cit. on p. 47).

[76] Andy Buckley et al. "General-purpose event generators for LHC physics". In: *Phys. Rept.* 504 (2011), pp. 145–233. DOI: 10.1016/j.physrep.2011.03.005. arXiv: 1101.2599 [hep-ph] (cit. on p. 47).

[77] Bryan R Webber. "Hadronization". In: hep-ph/9411384. CAVENDISH-HEP-94-17 (1994). URL: https://cds.cern.ch/record/272660 (cit. on p. 48).

[78] B. Andersson, G. Gustafson, and C. Peterson. "A semiclassical model for quark jet fragmentation". In: *Zeitschrift für Physik C Particles and Fields* 1.1 (Mar. 1, 1979), pp. 105–116. ISSN: 1431-5858. DOI: 10.1007/BF01450386. URL: https://doi.org/10.1007/BF01450386 (cit. on p. 48).

[79] B. R. Webber. "Fragmentation and hadronization". In: *Int. J. Mod. Phys.* A15S1 (2000). [eConfC990809,577(2000)], pp. 577–606. DOI: 10.1142/S0217751X00005334. arXiv: hep-ph/9912292 [hep-ph] (cit. on p. 48).

[80] Serguei Chatrchyan et al. "Jet and underlying event properties as a function of charged-particle multiplicity in proton?proton collisions at $\sqrt{s}$ = 7 TeV". In: *Eur. Phys. J.* C73.12 (2013), p. 2674. DOI: 10.1140/epjc/s10052-013-2674-5. arXiv: 1310.4554 [hep-ex] (cit. on p. 48).

[81] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "PYTHIA 6.4 Physics and Manual". In: *JHEP* 05 (2006), p. 026. DOI: 10.1088/1126-6708/2006/05/026. arXiv: hep-ph/0603175 [hep-ph] (cit. on pp. 49, 124).

[82] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "A Brief Introduction to PYTHIA 8.1". In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820 [hep-ph] (cit. on pp. 49, 124).

[83] J. Alwall et al. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations". In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph] (cit. on pp. 49, 124, 144).

[84] Johan Alwall et al. "Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions". In: *Eur. Phys. J.* C53 (2008), pp. 473–500. DOI: 10.1140/epjc/s10052-007-0490-5. arXiv: 0706.2569 [hep-ph] (cit. on p. 49).

[85] Rikkert Frederix and Stefano Frixione. "Merging meets matching in MC@NLO". In: *JHEP* 12 (2012), p. 061. DOI: 10.1007/JHEP12(2012)061. arXiv: 1209.6215 [hep-ph] (cit. on p. 49).

[86] Stefan Hoeche et al. "Matching parton showers and matrix elements". In: *HERA and the LHC: A Workshop on the implications of HERA for LHC physics: Proceedings Part A*. 2005, pp. 288–289. DOI: 10.5170/CERN-2005-014.288. arXiv: hep-ph/0602031 [hep-ph]. URL: https://inspirehep.net/record/709818/files/arXiv:hep-ph_0602031.pdf (cit. on p. 49).

[87] Fabio Maltoni. *Predictive Monte Carlo tools for LHC physics*. Academic Training Lecture Regular Programme. CERN, May 2012 (cit. on p. 50).

[88] Carlo Oleari. "The POWHEG-BOX". In: *Nucl. Phys. Proc. Suppl.* 205-206 (2010), pp. 36–41. DOI: 10.1016/j.nuclphysbps.2010.08.016. arXiv: 1007.3893 [hep-ph] (cit. on p. 49).

[89] Paolo Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146 [hep-ph] (cit. on p. 49).

[90] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph] (cit. on p. 50).

[91] S. Agostinelli et al. "GEANT4: A Simulation toolkit". In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8 (cit. on p. 50).

[92] Albert M Sirunyan et al. "Particle-flow reconstruction and global event description with the CMS detector". In: (2017). arXiv: 1706.04965 [physics.ins-det] (cit. on pp. 54, 56).

[93] W. Adam et al. "Track reconstruction in the CMS tracker". In: (2005) (cit. on p. 54).

[94] *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET*. Tech. rep. CMS-PAS-PFT-09-001. Geneva: CERN, Apr. 2009. URL: https://cds.cern.ch/record/1194487 (cit. on p. 54).

[95] W. Adam et al. "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC". In: *J. Phys. G* 31 (2005), N9. DOI: 10.1088/0954-3899/31/9/N01 (cit. on p. 56).

[96] G. L. Bayatian et al. "CMS physics: Technical design report". In: (2006) (cit. on p. 57).

[97] Andreas Hoecker et al. "TMVA: Toolkit for Multivariate Data Analysis". In: *PoS* ACAT (2007), p. 040. arXiv: physics/0703039 (cit. on p. 58).

[98] Philipp Pigard and Claude Charlot. "Electron studies and search for vector boson scattering in events with four leptons and two jets with the CMS detector at the LHC. Identification des lectrons et mise en vidence de la diffusion de bosons massifs dans les vnements quatre leptons et deux jets avec le dtecteur CMS auprs du LHC". Presented 2017. 2017. URL: http://cds.cern.ch/record/2290139 (cit. on p. 59).

[99] *Identification of b quark jets at the CMS Experiment in the LHC Run 2*. Tech. rep. CMS-PAS-BTV-15-001. Geneva: CERN, 2016. URL: https://cds.cern.ch/record/2138504 (cit. on pp. 59, 64).

[100] Henning Kirschenmann. *Jets at CMS and the determination of their energy scale*. CERN, July 2012. URL: http://cms.web.cern.ch/news/jets-cms-and-determination-their-energy-scale (cit. on p. 60).

[101] Gavin P. Salam. "Towards Jetography". In: *Eur. Phys. J.* C67 (2010), pp. 637–686. DOI: 10.1140/epjc/s10052-010-1314-6. arXiv: 0906.1833 [hep-ph] (cit. on pp. 59, 60).

[102] Vardan Khachatryan et al. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV". In: *JINST* 12.02 (2017), P02014. DOI: 10.1088/1748-0221/12/02/P02014. arXiv: 1607.03663 [hep-ex] (cit. on p. 61).

[103] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet User Manual". In: *Eur. Phys. J.* C72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph] (cit. on p. 61).

[104] Albert M Sirunyan et al. "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV". In: *JINST* 13.05 (2018), P05011. DOI: 10.1088/1748-0221/13/05/P05011. arXiv: 1712.07158 [physics.ins-det] (cit. on p. 63).

[105] Albert M Sirunyan et al. "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV". In: *Submitted to: JINST* (2017). arXiv: 1712.07158 [physics.ins-det] (cit. on pp. 63, 65).

[106] Serguei Chatrchyan et al. "Identification of b-quark jets with the CMS experiment". In: *JINST* 8 (2013), P04013. DOI: 10.1088/1748-0221/8/04/P04013. arXiv: 1211.4462 [hep-ex] (cit. on p. 64).

[107] Vardan Khachatryan et al. "Event generator tunes obtained from underlying event and multiparton scattering measurements". In: *Eur. Phys. J.* C76.3 (2016), p. 155. DOI: 10.1140/epjc/s10052-016-3988-x. arXiv: 1512.00815 [hep-ex] (cit. on p. 68).

[108] *Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV*. Tech. rep. CMS-PAS-PFT-10-002. Geneva: CERN, 2010. URL: https://cds.cern.ch/record/1279341 (cit. on p. 71).

[109] Vardan Khachatryan et al. "Search for new physics in same-sign dilepton events in proton?proton collisions at $\sqrt{s} = 13\,\text{TeV}$". In: *Eur. Phys. J.* C76.8 (2016), p. 439. DOI: 10.1140/epjc/s10052-016-4261-z. arXiv: 1605.03171 [hep-ex] (cit. on p. 79).

[110] CMS SUSY Group. "Computing the contamination from fakes in leptonic final states". In: AN-2010/261 (2010) (cit. on p. 80).

[111] A. Abulencia et al. "Measurements of inclusive W and Z cross sections in p anti-p collisions at s**(1/2) = 1.96-TeV". In: *J. Phys.* G34 (2007), pp. 2457–2544. DOI: 10.1088/0954-3899/34/12/001. arXiv: hep-ex/0508029 [hep-ex] (cit. on p. 91).

[112] *CMS Luminosity Measurements for the 2016 Data Taking Period*. Tech. rep. CMS-PAS-LUM-17-001. Geneva: CERN, 2017. URL: https://cds.cern.ch/record/2257069 (cit. on p. 102).

[113] *CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update*. Tech. rep. CMS-PAS-LUM-13-001. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1598864 (cit. on p. 102).

[114] Michiel Botje et al. "The PDF4LHC Working Group Interim Recommendations". In: (2011). arXiv: 1101.0538 [hep-ph] (cit. on p. 103).

[115] D. de Florian et al. "Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector". In: (2016). DOI: 10.23731/CYRM-2017-002. arXiv: 1610.07922 [hep-ph] (cit. on pp. 104, 110).

[116] *Procedure for the LHC Higgs boson search combination in Summer 2011*. Tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11. Geneva: CERN, Aug. 2011. URL: https://cds.cern.ch/record/1379837 (cit. on p. 105).

[117] F. James. "MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1". In: (1994) (cit. on p. 105).

[118] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *Annals Math. Statist.* 9.1 (1938), pp. 60–62. DOI: 10.1214/aoms/1177732360 (cit. on p. 106).

[119] L. Moneta et al. "The RooStats project". In: *Proceedings of the 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research. February 22-27, 2010, Jaipur, India.* 2010. arXiv: 1009.1003. URL: http://adsabs.harvard.edu/abs/2010acat.confE..57M (cit. on p. 106).

[120] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J.* C71 (2011). [Erratum: Eur. Phys. J.C73,2501(2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0, 10.1140/epjc/s10052-013-2501-z. arXiv: 1007.1727 [physics.data-an] (cit. on pp. 106, 107).

[121] A L Read. "Modified frequentist analysis of search results (the CL_s method)". In: CERN-OPEN-2000-205 (2000). URL: https://cds.cern.ch/record/451614 (cit. on p. 107).

[122] J Mans et al. *CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter*. Tech. rep. CERN-LHCC-2012-015. CMS-TDR-10. Sept. 2012. URL: https://cds.cern.ch/record/1481837 (cit. on p. 117).

[123] G Apollinari et al. *High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2116337 (cit. on p. 118).

[124] J Butler et al. *CMS Phase II Upgrade Scope Document*. Tech. rep. CERN-LHCC-2015-019. LHCC-G-165. Geneva: CERN, Sept. 2015. URL: https://cds.cern.ch/record/2055167 (cit. on p. 121).

[125] Li Qiang. *Vector boson scattering and fusion results from ATLAS and CMS*. URL: https://indico.cern.ch/event/442390/contributions/1096018/ (cit. on p. 123).

[126] Alessandro Ballestrero et al. "PHANTOM: A Monte Carlo event generator for six parton final states at high energy colliders". In: *Comput. Phys. Commun.* 180 (2009), pp. 401–417. DOI: 10.1016/j.cpc.2008.10.005. arXiv: 0801.3359 [hep-ph] (cit. on p. 123).

[127] J. de Favereau et al. "DELPHES 3, A modular framework for fast simulation of a generic collider experiment". In: *JHEP* 02 (2014), p. 057. DOI: 10.1007/JHEP02(2014)057. arXiv: 1307.6346 [hep-ex] (cit. on p. 124).

[128] Morad Aaboud et al. "Search for anomalous electroweak production of WW/WZ in association with a high-mass dijet system in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Phys. Rev.* D95.3 (2017), p. 032001. DOI: 10.1103/PhysRevD.95.032001. arXiv: 1609.05122 [hep-ex] (cit. on p. 145).